

Eötvös Loránd Tudományegyetem
Társadalomtudományi Kar

Németh Renáta

Oksági következtetés az empirikus szociológiai kutatásban

Habilitációs dolgozat

Budapest, 2014.

Köszönetnyilvánítás

Öt éve védtem az ELTE Társadalomtudományi Karán doktori disszertációm, s öt éve dolgozom itt főállású oktatóként. Habilitációs dolgozatomban az azóta végzett munkám oksággal kapcsolatos vonatkozásait gyűjtöttem egybe. Ezek a munkák nem valósultak volna meg a Kar támogató közege nélkül. Egykori témavezetőm, Rudas Tamás ma is legfontosabb kutatótársam; e munka utolsó készítése során is sok hasznos beszélgetést folytattunk. Tanszéki kollégám és szobatársam, Simon Dávid mindig inspiráló beszélgetőtárs. Fokasz Nikosz és a körülötte szerveződött Peripato kutatócsoport tagjaival az utóbbi években folytatott műhelybeszélgetésekből sokat merítettem; a dolgozat nagy részét is átbeszéltem velük. A tanulmány Big Data-vonatkozású fejezetét Csepeli György inspirálta, s néhány hivatkozását Dessewffy Tibortól kaptam. Vokó Zoltán révén ismertem meg a téma biostatistikai, kísérlettervezési vonatkozásait. A Kar fogadta be azt az oksággal kapcsolatos kutatómunkát is, melyet az MTA Bolyai János Kutatási Ösztöndíja támogat, ez az ösztöndíj tette lehetővé ennek a munkának az elkészültét is.

Köszönet értük.

Tartalomjegyzék

1. Bevezetés.....	4
2. Motiváció	4
3. Oksági következtetés az empirikus szociológiai kutatásban: alternatív fogalmi megközelítések	7
1.1. A probléma megközelítésének módja	7
1.2. Az okság hétköznapi fogalmának holdudvara	9
1.3. A társadalomtudományi megismerési módok tipológiája.....	10
1.4. Az okság mint metafizikai probléma	16
1.5. Az oksági következtetés négy típusa.....	20
1.5.1. Okság, mint robusztus összefüggés, avagy neo-humeiánus regularitás.....	20
1.5.2. Okság, mint a tényellentétes kijelentés teljesülése.....	23
1.5.3. Okság, mint hatást kiváltó manipuláció	25
1.5.4. A Rubin-Holland modell, mint a tényellentétes és a manipulációs modellt egyesítő statisztikai modell	27
1.5.5. Okság, mint mögöttes mechanizmus, vagy generatív folyamat.....	35
1.5.6. A négy megközelítés összevetése	39
1.6. Szociológiatörténeti vonatkozások.....	39
1.7. Mai tendenciák: Big Data – szükségtelen kauzalitás?	43
4. Megfigyeléses vizsgálatok a Rubin-Holland modell keretében: <i>confounding</i>	44
5. Grafikus modellek és kauzalitás.....	61
6. Befejezés	68
7. Irodalom	69

1. Bevezetés

Elfogadott álláspont, hogy a társadalomkutatás adatgyűjtési módszerei általában nem alkalmasak oksági következtetések levonására. Ennek ellenére az empirikus társadalomtudományi kutatás elsődleges célja nyilván nem csupán a társadalmi valóság leírása, hanem annak megértése oksági viszonyok feltárása révén.

Dolgozatom első felében (3. fejezet) arra a kérdésre keresek választ, hogy milyen (látens vagy explicit) fogalmi bázison működnek a mai társadalomkutatások az oksági következtetés szempontjából. Milyen „bizonyítékot” tartanak az egyes megközelítések szükségesnek/elegendőnek az oksági kapcsolatra történő következtetéshez? Röviden kitérek az okság megközelítésének a szociológia története során megfigyelhető változásaira és a mai tendenciákra is. A dolgozat ezen része nagyban épül friss vagy elbírálás alatt levő publikációimra (Németh, 2013, Németh, 2014, Németh, Rudas, 2014).

A dolgozat második felében (4. és 5. fejezet) ennek az áttekintésnek a hasznát próbálom megmutatni: a tényellentétes modell szempontjából tárgyalom a szociológiában fontos megfigyeléses vizsgálatok esetét, és az utóbbi években általam kutatott grafikus modelleket is elhelyezem az okság kontextusában. Ezek a fejezetek részben szintén a doktori disszertációm elkészítése óta megjelent munkáimra (Rudas, Bersgma, Németh, 2010, Németh, Rudas, 2013a, 2013b, Rudas, Németh, 2013, Németh, Rudas, 2014) épülnek.

2. Motiváció

„Amikor a jelenségeket a megfigyelő akarata szerint mesterségesen tudjuk létrehozni, akkor a szó szoros értelmében vett kísérleti módszerről van szó. Amikor viszont – épp ellenkezőleg – a tényeket nem mi hozzuk létre, és csak abban a formában tudjuk összevetni őket egymással, amelyben spontán módon kialakultak, akkor módszerünket közvetett kísérletnek vagy összehasonlító módszernek nevezzük. ...Minthogy... a szociológiai jelenségek kívül esnek a kutató hatáskörén, csupán az összehasonlító módszer felel meg a szociológiának.”

Durkheim, 1894¹

„És mégis: egyáltalán nem hiábavaló feltenni például azt a kérdést, hogy mi történt volna, ha Bismarck nem a háború mellett dönt. Ugyanis a valóság történelmi konstrukciójában éppen ennek a kérdésnek van meghatározó szerepe: a kérdésnek, miszerint a végtelen sok tényező együtteséből (mely tényezők mindegyikének éppen úgy kellett történnie, ahogy történt, és nem másként, ahhoz, hogy az adott következmény létrejöjjön) ennek az egyéni döntésnek mekkora kauzális jelentőséget tulajdoníthatunk; a kérdésnek, miszerint milyen helyet foglal el az adott

¹ A szociológia módszertani szabályai. In: Durkheim, Émile: *A társadalmi tények magyarázatához*. Bp.: Közgazdasági és Jogi Kiadó, 1978. pp 141.

*döntés a történet-vázban. Ha a történelem több akar lenni érdekes események és személyiségek krónikájánál, akkor ezt a célt csak ilyen kérdéseket feltevésével érheti el.”*²

Weber, 1906, Eduard Meyer könyvéről, melyben Meyer megválaszolhatatlannak és üresnek nevezi a Bismarck döntésére vonatkozó mi-történt-volna-ha kérdést

A két idézet jól mutatja, hogy a **kauzalitás problémájával már a szociológia megszületésekor szembesültek** a születésnél bábáskodók. A problémát elsősorban az jelentette, hogy a társadalomtudományok - a természettudományokkal szemben – tipikusan sajátos empirikus anyagból kell, hogy kiinduljanak. Ez az anyag egyedi események megfigyelése révén összegyűjtött tények összessége, mely események létrejöttére, létrejöttük körülményeire – a kísérletes természettudományokkal szemben – a megfigyelő nem lehet befolyással. Durkheim idézett gondolata a kísérletek és megfigyeléses vizsgálatok első, klasszikus szembeállítás. Ezután három évvel jelenteti meg Az öngyilkosság-ot, ahol megfigyeléses adatokból kiindulva, a megfigyelt együttjárások alapján alkot oksági magyarázatokat. Weber fent idézett írása pedig az első munkák egyike a tényellentétes (mi lett volna, ha nem...) állítások tudományos magyarázatban betöltött szerepével kapcsolatban. E két, karakteresen különböző okság-megközelítés, mint látni fogjuk, közvetlenül összekapcsolható az oksági következtetés társadalomtudományi megközelítésének egy-egy mai irányzatával.

Az okság problémája azonban nem csak a szociológia születésekor tartozott a legfontosabb, a szaktudományt definiáló kérdések közé. A kauzalitással foglalkozó mai szociológiai írások alapján úgy tűnik, **a problémának továbbra is nagy a tétje** - a szerzők gyakran hivatkoznak a tudomány - kauzális következtetések révén elérhető - kíváncsú előrehaladására. John Goldthorpe a „Megfigyelhető-e fejlődés, tudás-kumulálódás a szociológiában?” kérdésre adott válaszában (2005) többször kitér a kauzalitás kérdésére, miközben a szociológia lehetőségeit a természettudományokéval veti össze, a két terület lényegi különbségeit tagadva.³ Marini és Singer (1988) szintén a társadalomtudományok kíváncsú tudás-kumulálódásával hozza összefüggésbe a kauzalitás fogalmának tisztázását. Cikkükben a fogalom filozófiai és operacionális megközelítésének integrálását kísérik meg, a lezárásában kifejtve: „Meggyőződésünk, hogy mindez a társadalomtudományokban a kauzalitás bizonyíthatóságának szisztematikus evolúciójához, az új tudás jobban szervezett kumulatív fejlődéséhez vezet majd.” Az előrehaladás, a természettudományok szintjére való

² Saját fordítás, forrás: Weber, Max (1985 [1906]): Kritische Studien auf dem Gebiet der kulturwissenschaftlichen Logik. In: *Gesammelte Aufsätze zur Wissenschaftslehre*. 6. kiadás, Tübingen: Mohr, pp 267.

³ Goldthorpe tudás-kumulálódási kérdésére rimel a neves hálózatkutató, Nicholas A. Christakis The New York Times-beli, 2013-as provokatív publicisztikája. Idézet a cikkből: „A szociológusoknak a természettudósokhoz hasonlóan képesnek kellene lenni ilyen kijelentések megtételére: »Ezt a témát már elég nagy mélységben feldolgoztuk, így most más, izgalmasabb területekre térünk át.«” Érdekes megemlíteni, hogy a természettudományok felől érkező kihívás ma leginkább talán éppen a network-szociológiában érhető tetten.

felkapaszkodás régóta áhított reménye sokszor frusztráltan, kisebbségi érzésekkel telve fejeződik ki. Sørensen 1998-as cikkében a Goldthorpe-éhoz hasonló pozitivista alapállásból, az oksági magyarázatokban a pusztán statisztikai megoldásokon túl a társadalmi mechanizmusok megértésének szerepét hangsúlyozva fejti ki, hogy „a szociológia sosem lesz képes a tudományok hierarchiájában elfoglalt megalázó pozíciójából feljebb mozdulni, ha nem tudjuk a megbízható tudást és a hatékony elméleteket jobban integrálni. [...] A kvalitatív kutatás által létrehozott tudás sosem lesz annyira megbízható és általánosítható, hogy egy teljes tudományos elmélet kívánalmainak megfeleljen. [...] viszont [...] A kvantitatív szociológia jelenleg kevésbé épül az elméletre és kevésbé releváns az elmélet-fejlesztés számára, mint három évtizeddel ezelőtt.”

A kvantitatív empirikus szociológia kauzalitással kapcsolatos óvatossága a hazai szociológiában sem maradt reflektálatlan. Ez az óvatosság tartalmi kiüresedéshez vezethet – írja Némedi a szociológia jövőjéről írt nagyhatású cikkében (2000). A feltámasztott és a mai szociológiával szembeesülő Durkheim szájába adva teszi fel a kérdést: „Hol vannak a törvények, az oksági összefüggések, amelyek megállapítása a tudomány feladata? Mire való a tények gyűjtése, ha belőlük nem tudunk szilárd elméleteket felépíteni?”

A szociológia-oktatás módszertani oldalát a korrelációval kapcsolatban leginkább a tiltás jellemzi. A figyelmeztetés, miszerint az együttjárásból nem következtethetünk oksági kapcsolatra. A „*correlation does not imply causation*” jelmondat ismert közhely, a google 1.200.000 találatot ad rá, az első találat egy Wikipédia-oldal a mondatnak szentelve. A jelmondat ellenére a publikált empirikus tanulmányok jó része kizárólag korrelációs bizonyítékokkal érvel. Megfigyelhető stratégia, hogy a szerzők az elemző fejezetben a direkt oksági fogalmakat gondosan kerülő nyelvet használnak (kifejezések: korreláció, asszociáció, függő/független változó), míg a megbeszélésben és az absztraktban az eredményeket az oksági interpretáció felé tolják. Ez az eltolás néha direkten történik („az általunk azonosított faktoroknak ilyen és ilyen hatásai vannak”) néha indirekten („a szociálpolitikai intervencióknak az általunk feltárt rizikótényezőkre kell koncentrálnia” – sugallva, hogy a negatív hatások a beavatkozással megváltoztathatók). Más cikkek expliciten kifejtik, hogy csak asszociációkat, és nem oksági relációkat vizsgálnak, ám az eredmények így nagyjából érdektelenek lennének, ezért (őszinténül) arra számítanak, hogy az olvasó von le kauzális következtetéseket. Ezek a stratégiák - ha kicsit erős is a megfogalmazás - szabadulni próbálnak az elégséges oksági kijelentések igazolására vonatkozó kutatói felelősségtől⁴.

Hogy mindez ne tűnjön a kívülálló kritikájának, idéznék egy, a saját részvételemmel zajlott egészségfelmérés kutatási beszámolójából (Boros, Németh, Vitrai, 2002). A kutatási jelentés elemző fejezetei kauzálisan semleges „korrekt” nyelven vannak megfogalmazva, ám a megrendelő (Egészségügyi Minisztérium) miatt a jelentés „Az eredmények felhasználhatósága egészségpolitikai célokra” c. fejezettel bővült, ami oksági megfontolásokat kívánt. Itt írtuk: „A kutatás egyik legfontosabb tanulsága, hogy a funkcionalitást korlátozó

⁴ Gárdos Judit (2014) esettanulmányában szintén az asszociáció-elemzés és a kauzális konklúzió közti “hiányzó láncszemet” keresi. Egy konkrét előítélet-vizsgálati tanulmány szövegét elemezve ezt a láncszemet a kutatók implicit (a többségi társadalom előítéletességére vonatkozó) feltételezéseiben találja meg.

rosszabb egészségi állapot előfordulása magasabb azok között, akik gazdasági helyzetüket szubjektíve rosszabbnak érzik másokénál. Vélhetően minden, a társadalmi különbségek mérséklését szolgáló célzott beavatkozás jótékony hatással lehet az egészségi állapotban megmutatkozó egyenlőtlenségek csökkenésére is.” A jelen munkához vezető saját személyes motivációim között éppen ez a szakadék szerepelt: a szakadék, ami általában a kutatások statisztikai eredményei (szignifikáns paraméterbecslések) és a (közpolitikai beavatkozási célpontok azonosítására vonatkozó igényre választ adni igyekvő) bizonytalan kauzális konklúziók között húzódik.

Ugyanakkor szem előtt kell tartani, hogy az oksággal kapcsolatos óvatosságnak történeti oka van, mégpedig a kauzalitásnak a Karl Pearson-tól eredeztethető és több évtizedig érvényben levő kizárása az alkalmazott statisztikai diskurzusból (lásd a szociológiai történeti áttekintést az 1.6. fejezetben).

3. Oksági következtetés az empirikus szociológiai kutatásban: alternatív fogalmi megközelítések

1.1. A probléma megközelítésének módja

A következő fejezetekben arra keresek választ, milyen (látens vagy explicit) fogalmi bázison működnek a mai társadalomkutatások az oksági következtetés szempontjából. Milyen „bizonyítékot” tartanak az egyes megközelítések szükségesnek/elegendőnek az oksági kapcsolatra történő következtetéshez? E megközelítések között az okság általuk alkalmazott fogalma alapján teszek különbséget. Fontos megjegyezni, hogy a szociológiai megközelítéseknek ez a kategorizációja nem teljesen felel meg azok módszertani klasszifikációjának. Például mind a regresszió-elemzés, mind a párosítás (*matching*) a potenciális zavaró változókra történő kontrollálást, a feltételes hatások kimutatását szolgálja, de (mint alább látni fogjuk) az előbbi alkalmazói az okságra mint robusztus összefüggésre gondolnak, az utóbbi viszont inkább a tényellentétes megközelítéshez sorolható.

Specifikálnom kell, mit jelent a tanulmány címében az „oksági következtetés”. Ezt első megközelítésben példákkal szeretném megmutatni. Következtetés alatt nem logikai levezetést értek, hanem az általánosabb „*inference*” megfelelőjeként bármilyen, az okságra hivatkozó, a következőkhöz hasonló kérdésekre adott tudományos választ:

Mekkora munkaerőpiaci előnyt biztosít a diploma?

Hat-e a jövedelmi helyzet a politikai részvételre?

Diszkriminálnak-e a munkaadók a roma álláskeresőkkal szemben?

Eredményes volt-e az OFA munkáltatóknak kiírt 2009-es foglalkoztatottság-megőrző programja?

Meggyőződésem, hogy a példák felhozása általában, de a jelen tárgykör esetén különösen, nem csupán retorikai fogás, nem csupán illusztrálása a fő mondanivalónak, hanem éppen ellenkezőleg: a példák specifikálják a mondanivalót. Ezzel egyben le is szűkítik a bennfoglalt

problémák körét: a területet, az ontológiai kiindulást (mi lehet ok és okozat?), a kérdésfeltevés logikai típusát egyaránt. A kauzalitás filozófiai irodalmának egyik kiindulópontja, David Hume például leggyakrabban biliárdgolyók ütközésének egymás mozgására gyakorolt hatását hozta fel az ok és okozat példájaként, vagy John L. Mackie, a XX. század filozófiai kauzalitás-koncepciójának meghatározó alakja is egy ház kigyulladását és annak lehetséges okait elemezte sokat idézett példáiban. Ezek a technikai, fizikai jellegű példák lényegesen különböznek az előzőekben felsorolt kérdésektől – néhány faktorról leírható, kevésbé komplex rendszerbe ágyazott, időben és térben jól lehatárolt partikuláris eseményeket vizsgálnak, tudattal nem rendelkező szereplőkkel. Egy biliárdgolyó-példán alapuló megközelítésből kiindulva nem csak nehéz, de általában lehetetlen is alkalmazni az arra kidolgozott fogalmi apparátust pl. annak a kérdésnek a vizsgálatára, hogy mekkora munkaerőpiaci előnyt biztosít a diploma.

A példákból kitűnik, hogy elsősorban oksági magyarázatok érdekelnek és nem előrejelzések. Mégpedig statisztikai jellegű kauzális magyarázatok, hiszen a determinisztikus, törvényre hivatkozó magyarázat alkalmazhatósága a társadalomtudományokban (szemben a természettudományokkal) nyilván csekély. Mivel pontosan szerettem volna látni az engem érdeklő kérdéstípusoknak a helyét a társadalomtudományi megismerés általános tipológiájában, az 1.3. fejezetet ennek a kérdésnek szenteltem.

Az okság szociológiai megközelítései nem tárgyalhatók függetlenül az okság hétköznapi (kognitív pszichológiai) fogalmától illetve a filozófiában megjelent különböző elemzési módjaitól. Ezért az alábbiakban elsőként a kauzalitás alapvető hétköznapi „fogalmi holdudvarát” vázolom fel (1.2. fejezet), majd megkísérlem az általam központi problémaként felvetett, társadalomtudományi területekre vonatkozó oksági következtetés kérdését elhelyezni a kauzalitást involváló egyéb tudományos megismerés-módok között (1.3. fejezet). Ezután az okság, mint metafizikai probléma lényeges kérdéseit ismertetem (1.4. fejezet), végül az oksági következtetés kortárs szociológiai megközelítéseit tipizálom az általuk alkalmazott okság-fogalom alapján (). Megközelítésem szándékoltan elnagyolt és kevésbé technikai, hiszen célom nem volt sem az egyes pozíciók formalizálása, sem belső filozófiai koherenciájuk vizsgálata. Sokkal inkább az egyes pozíciók viszonya érdekel, és az, hogy mit nyújthatnak annak a „terepen dolgozó” szociológusnak, aki valamiféle kauzális elemzést kíván elvégezni.

A kérdéskörhöz való közeledésem módja eleve kizár néhány lehetséges kérdésfelvetést. Egy rövid, csak a szociológiai okság-fogalom vizsgálatához elengedhetetlen áttekintésen (1.4 fejezet) túl nem tervezem a filozófiai kontextus részletezését. Amellett, hogy ez a kérdés nem témám, kérdésfeltevésem szempontjából nem is tartom elsődlegesnek: úgy vélem, a szociológia okság-felfogása inkább a statisztikában gyökerezik, s e két terület, a filozófia és a statisztika XX. századi kauzalitás-megközelítése egymástól meglehetősen elzártan fejlődött egymástól.

A szociológiai megközelítések közül is szelektálni kívánok. A téma pontos pozicionálása érdekében vázlatosan jelzem csak az egyéb, alternatív megközelítéseket – ahogy a példaként felsorolt kérdéseimből kitűnik, elsősorban az oksági hatás meglétét és nagyságát kérdésként megfogalmazó, statisztikai alapú hatásvizsgálatok foglalkoztatnak. Az oksági következtetés

versengő kortárs megközelítéseinek elméleti és metodológiai dilemmáit egyaránt részletezem. Bizonyos csomópontokra a meghatározó Rubin-modellen keresztül kívánok rávilágítani, mert ezt olyan módszertani eredménynek gondolom, amely a kérdéskör gyakorlati vonatkozásában is termékenyítő lehet. Ennek alátámasztásaként az utolsó fejezetekben a szociológiában meghatározó megfigyeléses vizsgálatokra alkalmazom a Rubin-modell szemléletét és próbálom megmutatni ennek következményeit. Végül, saját munkáimra térve, az utóbbi években általam kutató statisztikai modellek, a grafikus modellek⁵ viszonyát tisztázom a tárgyalt oksági koncepciókhoz, s helyezem őket kontextusba.

1.2. Az okság hétköznapi fogalmának holdudvara

A tudományos fogalomhasználat áttekintése előtt érdemes röviden elidőzni annál a kérdésnél, hogy hogyan jelenik meg az okság a hétköznapi gondolkodásban és nyelvhasználatban, hiszen nyilván a tudományos fogalomhasználat számára is ez jelent kiindulópontot. A kognitív nyelvész Lakoff és Johnson (1999) a kérdésre adott válaszában - más filozófiai fogalmak mellett - az ok és okozat elemzését is az angol nyelvben velük kapcsolatban használt metaforákból kiindulva végzi. Módszerük hátterét megtestesült tudat (*embodied mind*) koncepciójuk adja: úgy vélik, hogy elvont fogalmaink mögött is szenzomotoros rendszerünk és érzelmeink állnak. Az okság fogalmára térve: kiindulópontjuk az, hogy az angol metaforák az egyes állapotokat térbeli helyzettel azonosítják (*to be in a deep depression, to be in love, far from safety, close to insanity*). Ezen állapotok megváltozása helyzetváltoztatásként jelenik meg a metaforákban (*went crazy, come out of depression, went over the edge*). S végül, mindezzel összhangban: az okozott hatások kikényszerített helyzetváltoztatásokkal azonosíthatók (*The news propelled the stock market to record heights. FDR's leadership brought the country out of depression. That experience pushed him over the edge.*). A nyelvészeti bizonyítékot alátámasztja az is, hogy a jelenség produktív: eredetileg absztrakt kauzális értelemmel nem bíró igék, mint a *hurl* segítségével új, de az anyanyelvi beszélő számára világos jelentésű, absztrakt okságot kifejező metaforák képezhetők (*A sudden drop in prices hurled the farm belt into chaos*). A beszélők tehát nem tudják leválasztani az eredeti mozgásos képet a kauzalitás absztrakt fogalmáról, tehát a kauzalitás fogalmához a kikényszerített mozgás szorosan hozzátartozik.

A fenti példák gyakorlatilag közvetlenül magyarázhatók, pl: állapot: *mély depresszióban van* / állapotváltozás: *mély depresszióba zuhant* / kikényszerített állapotváltoztatás: *mély depresszióba taszította*. Az angolhoz hasonlóan a magyarban is produktív a jelenség: a *depresszióba taszít* a google szerint gyakran használt kifejezés (1100 találat), a *depresszióba lök* kevésbé gyakori (100 találat), a *depresszióba nyom* gyakorlatilag sosem használatos (2 találat); az anyanyelvi beszélő számára mégis világos az utóbbiak jelentése is. Tehát úgy

⁵ A legfontosabb hivatkozások: Németh R, Rudas T: On the application of discrete marginal graphical models. Sociological Methodology 43:(1) pp. 70-100. (2013), Németh, R., Rudas T: Discrete Graphical Models in Social Mobility Research: A Comparative Analysis of American, Czechoslovakian and Hungarian Mobility before the Collapse of State Socialism. Bulletin de Methodologie Sociolo 118:(1) pp. 5-21. (2013) Rudas, T., Bergsma, W., Németh, R. (2010): Marginal log-linear parameterization of conditional independence models. Biometrika

tűnik, Lakoff és Johnson megfigyelése a magyarra is áll, vagyis az emberi gondolkodásra univerzálisan teljesül.

Ide tartozik Spellman és Mandel 1999-es írása is, amely inkább tisztán pszichológiai, mint nyelvészeti megközelítésben foglalkozik az okság hétköznapi értelmezésével. Tanulmányuk cáfolja azt az elterjedt elképzelést, hogy hétköznapi oksági döntéseinket elsősorban a tényellentétes (mi lett volna, ha nem...) állítások határoznák meg. Ehelyett úgy találják, hogy a kérdés komplexebb: ha a hétköznapi beszélők szokatlan vagy negatív következmények potenciális megelőzési lehetőségét keresik, akkor valóban tényellentétest használnak („ha nem szállok autóba aznap, nem gázolom el”), de ha események okát akarják megtalálni, akkor az *okozattal szokásosan együttjáró faktorokat* keresik.

Mindennek akkor látjuk majd igazi jelentőségét, ha a kauzális következtetések versengő tudományos megközelítéseit vesszük sorra. A tudományos megközelítés nyilván nem független a hétköznapi fogalomhasználatától. A kikényszerített mozgás fogalma pl. leginkább a manipulációs megközelítéssel lesz rokonítható (hatás = egy manipuláció megfigyelt eredménye). A manipulációs megközelítés alternatívái között pedig ott fog szerepelni a tényellentétes megközelítés és az együttjárásokat vizsgáló neo-humeiánus regularitás-alapú megközelítés is.

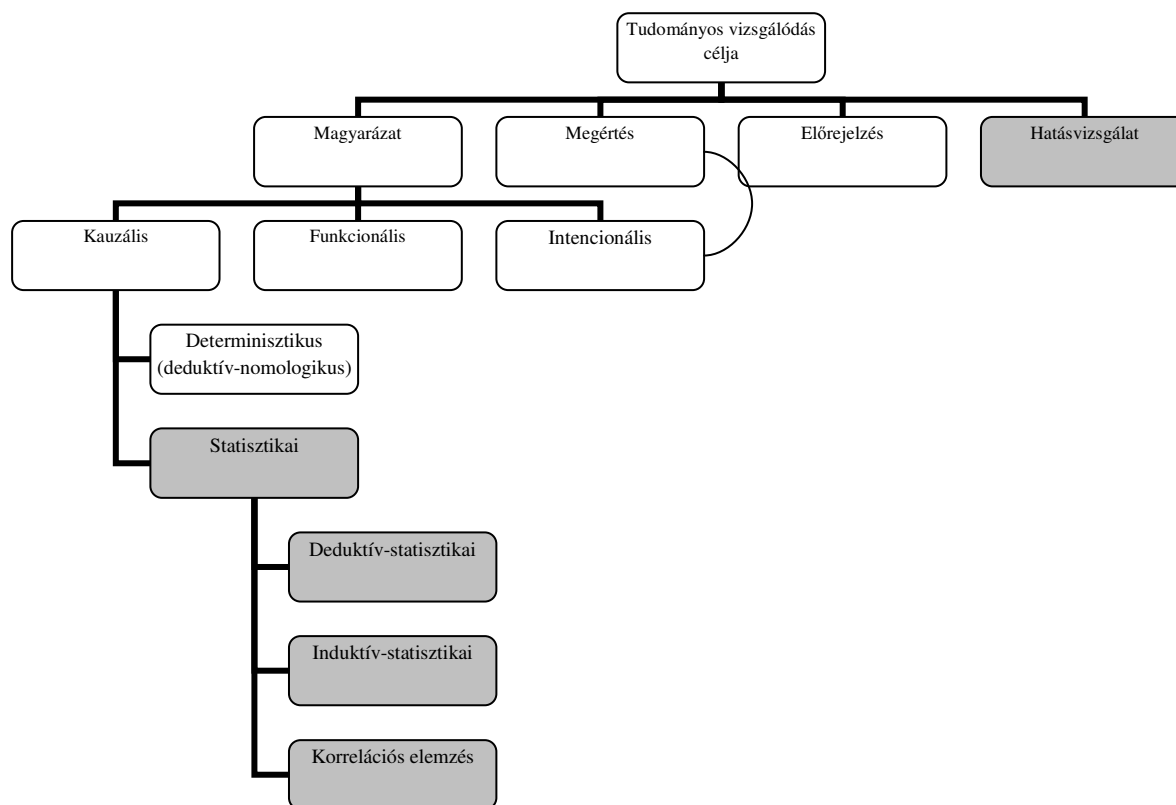
1.3. A társadalomtudományi megismerési módok tipológiája

A kauzalitás hétköznapi fogalomhasználata után térjünk rá a társadalomtudományi használatra. Első lépésben az érdekel, hogy az engem foglalkoztató, kauzalitásra hivatkozó és statisztikai eszközökkel megközelíthető problémák (lásd a korábban felsorolt típuskérdéseket: biztosít-e a diploma előnyt a munkaerőpiacon / az OFA munkáltatóknak kiírt 2009-es foglalkoztatottság-megőrző programjának volt-e pozitív hatása) hogyan helyezhetők el a társadalomtudományi megismerési módok között. Az alábbiakban ezért megkísérlem a tudományos megismerési formákat tipizálni, explicitté téve az általam vizsgálandó típusok megkülönböztető jegyeit.

A tudományos megismerést úgy definiálhatjuk, hogy nem csupán „leírását” adja a jelenségeknek, hanem valamely elmélet vagy modell által feldolgozza, strukturálja azokat. A megismerési módok tipizálását az 1. ábra szemlélteti, ahol az engem foglalkoztató típusokat (sárgítva) elkülönítettem a többitől. A tipizálás módja egyrészt a tudományos kérdésfeltevés logikája alapján (pl. okból indul ki: előrejelzés, okozatból indul ki: magyarázat), másrészt a tudományfilozófiai alapállás alapján (pl. magyarázat vs. megértés, funkcionális vs. intencionális magyarázat) történt.

Bár az ábra azt sugallhatja, hogy ezek egymást kizáró típusok, valójában nem mindig azok (lásd megértés és intencionális magyarázat érintkezését), a különbség inkább hangsúlybeli eltéréseken alapul. A tipizálás több szerző munkájára támaszkodik, ezeket saját céлом érdekében összefontam, részben átalakítottam és kiegészítettem. A kategorizáció célja nem egy szigorú és elméletileg teljesen koherens felosztás, a típusok inkább csak a jellemző különbségeket jelzik. Nem csak azért hangsúlyozom ezeket a különbségeket, hogy

megkönnyítsék a tárgyamhoz nem tartozó ágak lenyesését, hanem azért is, hogy jelezzék az általuk támasztott elméleti kihívásokat.



1. ábra A társadalomtudományi vizsgálódások tipológiája, sárgítva a számomra fontos, kauzalitásra hivatkozó és elsősorban statisztikai módszerekkel vizsgálható esetek

A magyarázat, megértés és előrejelzés klasszikus felosztása von Wright-tól ([1971],1987) jön. Determinisztikus esetben **magyarázat és előrejelzés** egyszerűen egymás időbeli megfordítottjai: ha fennáll a determinisztikus „ha X esemény bekövetkezik, akkor utána Y esemény is bekövetkezik” törvény, akkor X fennállásakor Y előrejelezhető, ahogyan Y előrejelzése X megfigyelése révén maga is potenciális magyarázatként szolgálhat Y-ra (Hempel, 1965). Ám általában ez nem áll. A társadalomtudományokban általában nem-determinisztikus rendszereket tételezünk fel – akár azért, mert alapvetően nem-determinisztikusnak gondoljuk a társadalmi valóságot, akár azért, mert a determinisztikus valóság teljes leírásához nincsen elegendő adatunk vagy megfelelő analitikus eszközünk.

A nem feltétlenül determinisztikus, általános esetet tekintve az **előrejelzés** általában egy modellre épülve történik, és ez a modell a legkisebb hibát adó előrejelzésre van optimalizálva. Ennek a modellnek nem kell feltétlenül tudnia a tartalmi szociológiai (pl. individuális döntési mechanizmusokra épülő) magyarázatról. Ebben a modellben ad-hoc változók is felhasználhatók, ha javítják az előrejelzést. Pl. választási előrejelzés adható exit-poll adatok alapján („Várhatóan megnyeri, mert a legtöbb megkérdezett rá szavazott”), de a parlamenti választások kimenetelének magyarázatához ez nem elegendő (helyette: „Megnyerte, mert az

ellenzéknek nem sikerült meggyőző programmal előállnia”). Ugyanis a **magyarázat** az előrejelzéssel szemben a ténylegesen működő oksági hatás modellezésére szolgál. Ahogyan az előrejelzés lehetősége nem implikálja magyarázat meglétét, úgy adekvát magyarázat is elképzelhető előrejelzésre való képesség nélkül⁶, erre jó példa Elster gyakran citált magyarázata az alkoholista szülőkkel kapcsolatban. Eszerint azt, hogy kiben mit vált ki az alkoholista környezet (maga is alkoholista, vagy inkább absztinens lesz), nem lehet előre megjósolni, de hogy ez a környezet nem marad hatástalan, az valószínű, és tulajdonképpen ez adja a magyarázatot. Egy másik, Elster mikroszintje után a makroszintet érintő példa: bár a társadalomtudományok kudarcot vallottak a kelet-európai kommunista rendszerek összeomlásának előrejelzése tekintetében, ennek ellenére utólag nagy számban születtek általános szükségszerűségekre épülő magyarázatok (az előrejelzés kudarcának lehetséges okairól lásd az American Sociological Association által kezdeményezett, az American Journal of Sociology hasábjain 1995 májusában közölt vitát). Ide tartozó példa Martikainen (1999) munkája is. Tanulmányában a munkanélküliségnek a mortalitásra gyakorolt kedvezőtlen hatását vizsgálja, valószínűsíthető kauzális kapcsolatot detektálva. Ugyanakkor megmutatja, hogy a rendelkezésre álló adatok és modellek alkalmatlanok lennének előrejelzésre, mert a növekvő gazdasági recesszióval párhuzamosan növekvő munkanélküliség hatása nem vezethető le a recessziót megelőző idők munkanélküliség okozta többlet-mortalitásából.

Míg a magyarázat ebben a kategorizációban törvényszerűségekre történő visszavezetést jelent, addig a **megértés** nem más, mint interpretáció. Megérteni az adott jelenséget annyit tesz, mint megpróbálni megtalálni azt a perspektívát, ahonnan nézve az érthető és elfogadható lesz, vagyis megpróbálni beilleszteni egy értelemösszefüggésbe. A magyarázattal szemben a **megértésnek** van pszichológiai (a jelenségbe történő beleérzés szükségessége), és intencionalisztikus eleme is. Értjük a cselekvést, ha értjük a cselekvő céljait (Max Weber), a szimbólumok, intézmények jelentését (kultúrantropológia).

A magyarázatokat Elster (1983) tudományfilozófiai megközelítésben kauzális, funkcionalista és intencionális magyarázat-sémákra osztja. A **kauzális magyarázat** (az ezen kategorizáció által adott szűk értelmében) törvényjellegű (*lawlike*) állításokra hivatkozik, melyek elégséges/szükséges okait adják meg az okozatoknak. Elster (1983) Hempel-re támaszkodva ezt a magyarázat-típust tovább specifikálja **determinisztikus** (deduktív-nomologikus) és **statisztikai** típusra. Az előbbi esetben megmagyarázzuk X-et, ha egy általános kauzális fedőtörvényre, *covering law*-ra történő hivatkozással az egyedileg megfigyelt, X-et időben megelőző eseményekből logikailag le tudjuk vezetni.

Az statisztikai típus esetén a törvény csak valószínűségi törvény, és a levezetés sem ad teljes, csak valószínűségi magyarázatot. Ennek a típusnak nagyobb a társadalomtudományi

⁶ Egy provokatív könyv erről Duncan J. Watts műve (Everything is Obvious - Once you know the answer. Crown Business: New York, 2011). Watts szerint a józan észre hivatkozó utólagos magyarázataink alapján tévesen azt gondoljuk, hogy értjük a társadalmi/gazdasági rendszerekben megfigyelt jelenségeket, pedig általában képtelenek vagyunk előrejelezni vagy irányítani azokat.

alkalmazhatósága, továbbá statisztikai eszközökkel vizsgálható, így tárgyamhoz tartozik, ezért az ábrán sárgítva szerepel. A valószínűség szerepet játszik az okság minden szociológiai megközelítésében. Mint látni fogjuk, a regularitás-elmélet szociológiai megvalósulása *statisztikai* együttjárásokra hivatkozik, ahogy a tényellentétes is a pozitív kimenet *valószínűségének* a „kezelés” hatására megvalósuló változásáról beszél. Ezzel szemben a filozófia, pl. Huoranszky (2001) az okság probabilisztikus felfogását a regularitás- ill. tényellentétes-elmélettől elkülönítve, önálló okság-meghatározásként definiálja. Itt hangsúlyoznám, hogy a kauzalitás probabilisztikus felfogása adott a szociológia esetében a statisztikának tekintélyt és munkát, és, mint alább olvasható, egyes szerzők szerint éppen az ad okot a kritikára, hogy a kauzalitás problémája sok esetben kizárólag statisztikai kérdéssé degradálódik.

Itt a determinisztikus és a statisztikai magyarázatot az általuk hivatkozott törvény típusa szerint különítettük el, de általánosabban, a vizsgált (fizikai vagy társadalmi) világ determinisztikusnak/nemdeterminisztikusnak történő feltételezése alapján is különbséget tehetünk az oksági magyarázatok között. Továbbá, statisztikai magyarázatokra jutunk akkor is, ha bár determinisztikus világot tételezünk fel, de a világot teljesen meg nem ismerhetőnek tartjuk. Erről bővebben lásd az 1.4. fejezetet.

Elster (1983) Hempel (1965) követve a statisztikai magyarázaton belül további kategóriákat különít el. Bár e kategorizáció számomra kissé önkényesnek tűnik, nem teljesen világos határvonalakkal, mégis érdemes röviden kitérni rá, hogy lássuk, milyen példák tartozhatnak a statisztikai magyarázathoz. Elster tehát megkülönböztet induktív-statisztikai és deduktív-statisztikai típust, továbbá, Hempel eredeti tipológiáját kiegészítve, korrelációs elemzést. A **deduktív-statisztikai** típus nem egyedi esettel, hanem általános mintázattal foglalkozik, egy sztochasztikus rendszert feltételezve, s a rendszert leíró statisztikai törvényszerűség(ek)ből (mint premisszákból) vezeti le ezt a statisztikai regularitást. Elster példáját véve adott iskolázottsági mobilitási mintázat magyarázatakor feltételezhetjük, hogy az intragenerációs mobilitás elsőrendű Markov-folyamatot követ, vagyis az iskolázottsági szintek közötti átmenet valószínűsége csak a pillanatnyi iskolázottság függvénye. Elster hangsúlyozza, hogy az ilyen modell bár mikroszintre visszavezetve magyarázza a makroszintű iskolázottsági mintázatot, nem ad elméleti indokot a mikroszintű mozzanatra (itt: a karrierutak közti választásra). Talán érdemes megjegyezni, hogy ugyanez áll a legtöbb magyarázó statisztikai modellre, pl. az adatokra illesztett lineáris regressziós egyenes általában csupán az adatokból kikövetkeztetett statisztikai regularitás, nem ad magyarázatot a lineáris összefüggés miértjére. Freedman (2010) szavaival: csupán görbeillesztés, a mögöttes adatgeneráló mechanizmus alátámasztása nélkül. Az **induktív-statisztikai** típus nem mintázatokat, hanem partikuláris eseményeket próbál megmagyarázni egy statisztikai törvényszerűségre/modellre hivatkozva (pl.: az adott egyén azért mozgott felfelé a mobilitási pályán, mert magas átmenetvalószínűség köti össze múltbeli és jelenlegi iskolázottsági szintjét). Hempel definíciójában az induktív és a deduktív statisztikai magyarázat közötti fő különbség az, hogy az előbbi individuális eseményre vonatkozik, míg az utóbbi statisztikai szabályosságra. Az utóbbi esetben a explanandum ténylegesen levezethető az explanansból, míg az előbbi esetben csak azt mondhatjuk, hogy nagy valószínűségű a bekövetkezése. Elster példáján

egyértelműnek tűnik az induktív és deduktív statisztikai magyarázat megkülönböztetése a társadalomtudományokra vonatkoztatva, de más példákon kevésbé egyértelmű az (egyetlen személyre vonatkozó) individuális eset és a (nagyobb mintaelemszámra épülő) mintázat megkülönböztetése. Elster az adott populáció megfigyelt mobilitási mintázatát (nyilván a nagy esetszám miatt) pl. nem individuális esetnek tartja, de pl. partikuláris esetnek minősül-e vagy már mintázatot alkot az OFA programjában résztvevő vagy abból kizárt cégek üzleti eredménye? Bár e kérdésnek a statisztikai megközelítés szempontjából lehet relevanciája, tárgyunkhoz nem tartozik szorosan.

Elster harmadik statisztikai magyarázat-típusa, a Hempelnél nem szereplő **korrelációs elemzés** szerint X magyarázza Y-t, ha erős korrelációt mutatnak. Itt nyilván fellép a hamis korreláció (*spurious correlation*) veszélye, vagyis ez a korreláció (vagy annak egy része) adódhat csupán X-nek és Y-nak egy harmadik változóval, Z-vel való kapcsolatából (Z-nek az X és Y közötti hatásmechanizmusban elfoglalt lehetséges helyéről részletesebben lásd az 0. fejezetet). Az 1.7 fejezet egy példájával: legyen X az adott földrajzi régió házassági arányszáma, Y pedig a régióban írt tweetek átlagos boldogsági szintje (a szöveg pozitív tartalmú kifejezéseinek arányával mérve). A köztük levő erős korreláció nem feltétlenül a házasság mellett szóló érv („a házasság boldoggá tesz”). Adódhat abból, hogy a tradicionálisabb államokban hagyományosan magasabb a házassági hajlandóság, ugyanakkor kulturális sajátossággént az írott szövegek emocionális töltöttsége is. Ennek a problémának a kivédésére a korrelációs elemzés újabb, potenciális zavaró változók bevonásával számolt parciális korrelációkkal (vagyis a korrigált, direkt kapcsolatok erősségének kiszámításával) próbálja megerősíteni a kauzalitás feltételezését. Elster ebbe a típusba sorolja a szorosan vett korrelációanalízisen kívül annak általánosításait (pl. útelelmzést) is.

Véleményem szerint nem teljesen határozott a deduktív-statisztikai és a korrelációs elemzésre támaszkodó magyarázat közötti különbség. Pl. a direkt és indirekt hatások, a hamis korreláció fogalma (bár a mérési szint miatt más asszociációs mérőszámra támaszkodva ugyan, de) a legtöbb többváltozós módszer, így az Elster példájában szereplő mobilitási Markov-folyamatmodelleknél is meghatározható. Elster aszerint tesz különbséget a két magyarázat-típus között, hogy a deduktív-statisztikai elemzés magában foglalja a mikroszintre történő visszavezetést, de pl. a lineáris regresszió esete mutatja, hogy ennek eldönthetősége kétséges: a lineáris regresszió a korrelációs elemzés egyfajta általánosítása, ugyanakkor tekinthetjük a változók közötti kapcsolat mikroszintre történő visszavezetésének is, ahol az egyéni jellemzőt (függő változót) más egyéni jellemzők lineáris függvényével leírható módon befolyásolnak. Elster nyilván azért egészítette ki Hempel sémáját a harmadik típussal, azért volt számára fontos e két típus elkülönítése, mert a magyarázatmódba a társadalmi mechanizmusokat is be kívánta emelni. A statisztikai együttjárások megtalálásán túl az együttjárást létrehozó, mikroszinten működő társadalmi mechanizmusokra történő hivatkozást tartotta fontosnak a magyarázat megalkotásában. Tulajdonképpen a mechanizmusokra épülő új okság-fogalmat használt, és ennek kereste a helyét Hempel rendszerében. A mechanizmusokra épülő oksági következtetést egyéb alternatíváival együtt a következő fejezetben tárgyalom majd részletesebben.

A tárgyunkhoz kevésbé tartozó másik két magyarázat-típusra térve röviden: a **funkcionális** magyarázat tipikusan pl. a kulturális antropológiában használatos; a társadalmi jelenségeket az általuk betöltött funkcióval magyarázza. A cselekvések és jelenségek a társadalom rendezett működéséhez és fennmaradásához való hozzájárulásukkal nyernek magyarázatot. Elster maga erősen megkérdőjelezte ennek a típusú magyarázatnak a társadalomtudományi alkalmazhatóságát, kifogásolva, hogy az emberek saját motivációi és ezekből fakadó döntések helyett olyan tényezőkkel magyarázunk jelenségeket, melyekről a résztvevők nem tudnak, vagy nem szándékoznak elérni. Elster szerint ezzel szemben az **intencionális** magyarázatnak van elsődleges szerepe a társadalomtudományokban, ez az a megkülönböztető jegy, ami elhatárolja őket a természettudományoktól. Ez a típus a társadalmi cselekvések magyarázatát a jövőre, azok céljára hivatkozva végzi, az emberi cselekvések mögött rejlő intenciók megértésével. A megértés történhet objektív vagy szubjektív perspektívából (Szántó, 1999). Amikor az objektív racionálisra hivatkozunk, akkor valamely kritérium alapján nyilvánítjuk helyesnek a cselekvést. A racionális választások sémája e magyarázat egyik fontos alelete, ilyenkor a várható haszon szempontjából optimális cselekvést nyilváníthatjuk racionálisnak. A szubjektíve racionális cselekvés magyarázatakor ezzel szemben a cselekvő perspektívájából, az ő érdekeit figyelembe véve keresünk magyarázatot - Weber szóhasználatában ez a motivációszerű megértés. A szubjektív és objektív racionalitás kettősége is állhat amögött, hogy az intencionális magyarázat, mint magyarázat-típus értelmezése erősen függ a tudományfilozófiai beállítódástól. Míg pl. az analitikus szociológus Hedström (2010) szerint már az „intencionális magyarázat” elnevezés is félrevezető, hiszen ez is egyfajta kauzális magyarázat (hiszen az egyéni cselekvések intencionális magyarázata mikromechanizmusokra visszavezetve történik), addig e pozitivista megközelítés ellenzői szerint az intencionális magyarázat nem kauzális jellegű, sőt nem is nomotetikus (nem hivatkozhat törvényjellegű állításokra), mert közelebb áll a megértéshez, mint a magyarázathoz (Bransen, 2001). Ezért kötöttem össze az ábrán az intencionális magyarázatot a megértéssel, jelezve a határvonal elmosódottságát.

A tanulmányom elején felsorolt három típuskérdésből kettő (biztosít-e a diploma előnyt a munkaerőpiacon, ill. az OFA munkáltatóknak kiírt 2009-es foglalkoztatottság-megőrző programjának volt-e pozitív hatása) nem sorolható be az eddig tárgyalt elfogadott kategóriákba. Ezeket a kérdéseket a hatásvizsgálatok közé sorolom. **Hatásvizsgálat** alatt itt oksági hatások bizonyítékon alapuló kimutatását és mérését értem. (Néha ez a kifejezés szűkebb értelemben használatos, közpolitikai intervenciók hatásának kimérésére.) A hatásvizsgálatot azért helyeztem negyedik fő típusként a magyarázat, megértés és előrejelzés klasszikus hármasa mellé, mert logikailag több szempontból is lényegesen különbözik azoktól. (1) Itt a magyarázattal és a megértéssel szemben az okból indulunk ki, nem a hatásból. (2) A kísérletesen igazolt hatásvizsgálati eredmény nem függ aktuális tudásunktól, ahogy Holland (1986) fogalmaz: „old, replicable experiments never die”. Ezzel szemben a magyarázat vagy az előrejelzés új ismeretek révén a jövőben mindig módosulhat. (3) A másik három típussal szemben okok leszűkített halmazán végezzük az elemzést, általában egy kitüntetett (*treatment*) és egy viszonyítási alapként szolgáló (*control*) ok hatását vetve össze. A diploma hatásaként szerzett munkaerőpiaci előny kérdésénél a kontroll a diploma hiánya,

az OFA programjának hatására vonatkozó kérdés esetén pedig a programban való rész nem vétel (vagy a pályázat elvesztése).

A hatásvizsgálatok és az előrejelzés viszonyával kapcsolatban érdemes megemlíteni a közgazdász Heckmant (2001), aki a ténylegesen implementált közpolitikai programok értékelése (mint hatásvizsgálat) mellett fontosnak nevezi ezek új környezetbe való (modell-alapú) extrapolációját és a programok hatásának ezen új környezetben való előrejelzését.

1.4. Az okság mint metafizikai probléma

Az előző fejezetben a társadalomtudományi megismerés általános módjait tipizáltam, a kauzalitásra hivatkozó és statisztikai eszközökkel vizsgálható módokat kiemelve. A tipológiában a kauzális magyarázat szűk, Elster által adott értelmében szerepelt: e kategória meghatározó jegye a valamely törvényjellegű állításra történő hivatkozás volt. A 1.5. fejezetben egy, az előzőtől részben független tipológiát ismertetek, mely már nem általában a tudományos megismerésnek, hanem a tudományos oksági következtetésnek (*causal inference*⁷) a statisztika által érintett területeit vizsgálja. Fontos látni majd, hogy hogyan kapcsolódnak ezek a kategóriák a fenti tipizáláshoz, mely területeket fedik le, s melyek maradnak lefedetlenek.

Míg a fenti tipológia elsősorban logikai kiindulású volt, a jelen tipológia magának a kauzalitás fogalmának az eltérő meghatározásaiból indul ki. Ezért itt a kauzalitás fogalmának a fenti, körvonalazatlan meghatározása már nem lesz elegendő. A pontosabb meghatározáshoz viszont elengedhetetlen a filozófiai kontextus rövid, csak a szociológiai alkalmazás szempontjából releváns pontokra koncentráló ismertetése (nem térek ki például azokra - a téma filozófiai irodalmának jelentős részét adó - analitikus érvekre, amelyek egyik vagy másik megközelítés elégtelenségét támasztják alá logikai szempontból). A filozófiai nézőpont bevonásakor meg kell említeni, hogy az oksági következtetés következő fejezetben ismertetendő megközelítései elsősorban a statisztikában gyökereznek. Az okság filozófiai és statisztikai elemzése viszont nagy távolságot tartott egymástól. Ahogyan pl. James Woodward 2003-as, *Making things happen: a theory of causal explanation* c. könyvében említi, a David Lewis és társai által a '70-es évektől kidolgozott filozófiai tényellentétes megközelítést a tényellentétes vele időben együtt fejlődő statisztikai tradíciója gyakorlatilag érintetlenül hagyta. És fordítva is: Holland (1986) statisztikai cikkéhez készült filozófiai jellegű hozzászólásokra reagálva bevallja, hogy nem ismerte fel a tényellentétes filozófiai megközelítésének saját modelljükkal meglevő párhuzamát. Ennek oka a megközelítések eltérő motivációjából eredhet: statisztikai nézőpontból randomizált kísérlettel határozható leginkább az oksági hatás, ugyanakkor ez a filozófia számára kevésbé lehet elfogadható, hiszen az okság megértése így egy aktív beavatkozást, egy aktív emberi tényezőt helyez a fókuszba.

⁷ A (kauzális) *következtetés* kifejezést itt nem a szűk logikai értelmében (premisszákból történő levezetés), hanem általános, az angol *inference*-re hivatkozó jelentésben használom: kauzális következtetés valamely kauzális állítás létrehozása.

Két általános szemponttal kezdeném, ami alapján az okság megközelítéseit vizsgálni érdemes. Az egyik szempont az okság természetére vonatkozó ontológiai kérdésekre, a másik szempont az episztemológiai kérdésekre adott válaszukon alapszik. Az **ontológiai kérdés** egyrészt annak a kérdése, hogy mi lehet ok és mi lehet okozat. Ez a szempont részben már fent is előkerül, pl. annak tárgyalásakor, hogy az okozat egy adott partikuláris jelenség (induktív-statisztikai magyarázat), vagy inkább valamely jelenség-együttes, mintázat (deduktív-statisztikai magyarázat). Kérdés lehet az is, hogy az okozatok események (térben/ időben lokalizálva, lásd az OFA 2009-es programja), tények (nem lokalizált: az OFA 2009-es programjában részt vett X cég növelte a bevételét), változók vagy (pl. társadalmi) struktúrák? Az okság klasszikus szerzője, Hume pl. a tárgyak (*objects*) mozgása között fennálló oksági kapcsolatot elemezte, de nem vizsgálta, milyen jelenségek tartozhatnak még ide. A szociológiában gyökeresen eltérő válaszokat találhatunk arra e kérdésre (lásd holizmus vs. individualizmus, realizmus vs. nominalizmus). Durkheim a szociológia, mint tudomány megszületése pillanatában deklarálta a szociológia sajátos tárgyainak, a társadalmi tényeknek az egyes egyénektől független létezését, és azt, hogy ezek magyarázatakor az okoknak is hasonló társadalmi tényeknek kell lennie (lásd: A társadalmi tények magyarázatához). Weber ezzel szemben nem ismerte el a kollektív cselekvők, intézmények reális létezését, ezek helyett az egyéni cselekvéseket, mint megmagyarázandó okozatokat helyezte szociológiája központjába, és okokként ezek szubjektív értelmét kereste (Gazdaság és társadalom, I: Módszertani alapok).

Ontológiai az a kérdés is, hogy ok és okozat között milyen jellegű kapcsolat állhat fenn. Más jellegű-e az oksági kapcsolat, ha biliárdgolyók ütközésére vonatkozik, mint ha társadalmi jelenségekre? Max Weber a kérdésre adott (a cselekvésnek értelmet tulajdonító cselekvőkre hivatkozó) válasza pl. közismerten igenlő. Fent hivatkozott munkájában azt állítja, hogy a szociológia tárgyai értelemmel bíró jelenségek, ezért a szociológiában léteznek ugyan általános valószínűségi törvények, de ezek a törvények az értelmet a cselekvésnek tulajdonító cselekvőtől nem lehetnek függetlenek. Egy tipikusnak mondott cselekvés korrekt kauzális interpretációja akkor valósul meg, ha nem csak a cselekvés tényleges előfordulásának várhatósága, vagyis a valószínűsége van korrekt módon megadva, hanem a cselekvés jelentése is adekvát módon van megragadva. Az előbbi inkább módszertani, az utóbbi tartalmi kérdés.

További kérdés, hogy milyen szerepe van az emberi ágensnek, az emberi pszichének az oksági kapcsolatban. Hume pl. az oksági kapcsolatot teljes egészében pszichológiai módon értelmezte, az emberi elme megszokáson alapuló hajlamaként (*propensity*) definiálva azt (Huoranszki, 2001).⁸

⁸ Más indokokra támaszkodva hasonló megállapításra jut a kauzalitást statisztikai modellekkel megközelítő Nobel-díjas közgazdász, Heckman (2005) is. „*Causality is in the mind*” – fogalmaz, azzal indokolva kijelentését, hogy kauzális modelljei csupán mentális konstrukciók, és a kauzális hatást is úgy definiálja, mint bizonyos változóknak a modellben történő hipotetikusan manipulálását. Heckman koncepciójában a kauzális hatás egy strukturális egyenletek modellezés paraméterbecslése, szemben a Rubin-Holland tradícióval, ahol a hatás a valóságban is megfigyelhető, megmérhető jelenség: egy kontrollált kísérletben megfigyelt különbség a kezelt és a kontroll csoport között. A két nézőpont (munkánéven legyen idealista és realista) szembenállása Heckman-nak

Ontológiai jellegű továbbá az a kérdés is, hogy az oksági kapcsolatban milyen szerepe van az időnek. Hume és majdnem minden más megközelítés is kritériumként szabja, hogy az ok mindig megelőzi az okozatot. Végül, ahogyan már láttuk, az oksági kapcsolat természetének determinisztikus vagy probabilisztikus jellege is lehet kérdés.

Eddig az okság természetével kapcsolatos ontológiai kérdéseket említettem. Az oksággal kapcsolatos másik szempont **episztemológiai**, és az empirikus szociológia számára elsődleges jelentőségű: hogyan ismerhetjük fel / bizonyíthatjuk be, hogy kauzalitás áll a megfigyelt jelenségek hátterében? Durkheim pl. fent hivatkozott munkájában az oksági viszonyoknak a megfigyelhetőségét, bizonyíthatóságát, és a bizonyítás alapvető módszertani szabályait (lényegében a korrelációs elemzést) is deklarálta.

A főbb ontológiai és episztemológiai szempontok említése után a továbbiakban ezen szempontok alapján ismertetném röviden az oksági viszonynak a filozófiában megjelenő jellegzetes elemzési módjait.

Említettem, hogy Hume az oksági kapcsolatot nem tartotta megfigyelhető empirikus kapcsolatnak, hanem az emberi elme egy hajlamaként értelmezte azt. Célja azon tapasztalati kritériumok megtalálása volt, melyek teljesülésekor hajlamosak vagyunk kauzalitást detektálni. Ezzel tulajdonképpen egyszerre válaszolta meg az okság ontológiai és episztemológiai kérdését. Az okságnak ezt a meghatározását a szociológiai okság-megközelítések egyike sem fogadná el, ennek ellenére Hume erős hatást gyakorolt a szociológiai gyakorlat okság-megközelítésére: kritériumai közül az egyik, az állandó együttjárás (*constant conjunction*) feltétele adja a következő fejezetben ismertetendő robusztos összefüggés-megközelítés lényegét (Hume másik két kritériuma is része az utóbbi megközelítésnek, ezek szerint az ok időben megelőzi az okozatot ill. ok és okozat térben/időben érintkeznek). Sokan ezt az együttjárás-kritériumot tekintették a Hume-i kritériumok legfontosabbikának, ezért nevezik a Hume-i indíttatású kauzalitás-megközelítéseket szabályszerűség- vagy **regularitás-elméleteknek** (Huoranszki, 2001). Vegyük azonban észre, hogy míg az eredeti Hume-i kritérium a tapasztalásra vonatkozott (együttjárást tapasztal a megfigyelő), addig az újabb megközelítések a tényleges együttjárásra hivatkoznak, az emberi psziché közbeiktatása nélkül.

Itt kell megemlíteni azokat a neohumeiánus filozófusokat (köztük a korábban már hivatkozott Hempelt), akik a regularitás-feltétel kiterjesztésével törvényjellegű állítások megfogalmazásával igyekeztek túllépni a véletlenszerű, esetleges regularitásokon. A modern hume-iánus felfogás tehát szabályszerű együttjárásokra hivatkozik, olyan kauzális törvényekre, melyek kimondják: mindig, amikor az ok fennáll, az okozat is bekövetkezik. Az okok ekképpen történő szükséges/elégséges feltételekként való azonosítása több logikai jellegű problémát okoz, ezek egyik megoldási javaslata John Mackie-hez fűződik (Brady,

Michael E. Sobel-lel folytatott vitájában fogalmazódott meg legmarkánsabban (2005, *Sociological Methodology*). Heckman legfőbb érve az, hogy a Sobel által preferált közvetlen összevetésnek a kísérleti és kezelt csoport érvényes összevethetőségére vonatkozóan sok olyan előfeltevése van, melyet sosem tesztelnek. Sobel realista nézőpontját (Holland és Cliff mellett) Freedman (2010) is osztja, aki általában is óv a statisztikai modellek uralmától.

2009). Mackie az okot un. INUS-feltételként írja le, ez az *Insufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result* kifejezés rövidítése. Az angol elnevezés arra utal, hogy az oknak nem kell feltétlenül szükségesnek lennie az okozat bekövetkezéséhez, továbbá az okok nem önmagukban állnak, hanem egyes (az okozat bekövetkezéséhez önállóan elégséges, de mivel nem egyedüli, ezért nem szükséges) ok-halmazok részei, s mint ilyenek, a saját ok-halmazuk működéséhez szükségesek, de a többi halmaz-elem nélkül nem elégségesek. Ahogy Huoranszki (2001) megjegyzi, Mackie előnye a klasszikus regularitás-elmélettel szemben az, hogy lehetőséget ad az ok(ok) azonosítására egy elégséges feltételrendszer elkülönítése által. A klasszikus elméletnek ugyanis nincs válasza arra, hogy a Hume-i három kritériumot teljesítő potenciális okok közül miért éppen egy bizonyos jelenséget tekintünk oknak, s miért nem pl. egy azt megelőző jelenséget. A szociológia számára Mackie rendszerének jelentősége a korábbi szűk ok-felfogás kitérítésében áll: megengedi több elégséges ok jelenlétét, megengedi azt, hogy egyazon ok több tényezőtől álljon (Mackie példájánál maradva a gyufa és az oxigén együtt alkotják a tűz okát), és megemlíti az okok közötti interakció lehetőségét. Az utóbbi pont kapcsán Brady (2009) számszerű példán mutatja be, hogy determinisztikus összefüggés probabilisztikusnak tűnhet akkor, ha a valóságban interakcióban működő okokat egymástól függetlennek tételezünk fel.

A **tényellentétesre** (vagy kontrafaktuális, *contrafactual*) hivatkozó megközelítés a regularitás-elmélet alternatívája. Míg az utóbbi regularitások, szabályszerűségek elemzését teszi lehetővé, addig az előbbi konkrét esetek, pl. történelmi események vizsgálatára alkalmas. Weber-nek a bevezetőben idézett kérdése jó példa a tényellentétesre támaszkodó okra: Bismarck döntése akkor oka a háborúnak, ha a háború nem tört volna ki Bismarck más irányú döntése esetén. David Lewis, a tényellentétes filozófiai elméletének kidolgozója szerint a tényellentétes „ha az ok nem állt volna fenn, akkor az okozat nem következett volna be” kijelentés igazsága a legközelebbi lehetséges világgal ellenőrizhető, amit úgy definiálhatunk, mint a miénkhöz leginkább hasonlót azon világok közül, ahol az ok nem állt fenn.

Míg a hume-iánus és a tényellentétes megközelítés szükségszerű okozat-bekövetkezéstről beszél, addig az okság **probabilisztikus felfogása** csak valószínűsíti azt. Meg kell különböztetni azt az esetet, amikor a bekövetkezésben azért nem vagyunk biztosak, mert bár determinisztikus oksági viszonyokat tételezünk fel, de nem ismerjük eléggé az összes fennálló tényezőt (ez a valószínűség szubjektív értelmezése), attól az esettől, amikor magát a vizsgálat tárgyát (mint a kvantummechanikában a világ mikrofizikai szerkezetét) tartjuk indeterminisztikusnak (Huoranszky, 2001). Az utóbbi, indeterminisztikus elméletet követve a valószínűséget nem feltétlenül kell szubjektív módon értelmezni ahhoz, hogy csak valószínűsített okozatokat tételezzünk fel: még ha minden körülményt pontosan ismerünk is, az indeterminisztikus fizikai világban az okozat bekövetkezése nem szükségszerű. A valószínűség objektív értelmezése szerint A és B esemény között akkor áll fenn oksági viszony, ha B-nek az A-ra vett feltételes valószínűsége nagyobb, mint az A fenn nem állása esetén vett feltételes valószínűség. Vélhetően a szociológusok egyetértenek velem abban,

hogy a társadalmi folyamatok és relációk többsége indeterminisztikus⁹ – ezért a kauzalitás probabilisztikus felfogásának a következőkben fontos szerepe lesz. Fontos megjegyezni, hogy míg a filozófiai megközelítésű Huoranszky (2001) ezt a felfogást a regularitás- ill. tényellentétes-elmélettel összeegyeztethetetlennek tartja, ezért egy önálló harmadik filozófiai okság-meghatározásként különíti el, a kauzális következtetés szociológiai megközelítéseit tekintve ez kevésbé áll. Mint látni fogjuk, a szociológiai megközelítésekben a valószínűség az előbbi két megközelítésben is szerepet játszik: a regularitás-elmélet szociológiai megvalósulása *statisztikai* együttjárásokra hivatkozik, ahogy a tényellentétes is a pozitív kimenet *valószínűségének* kezelés (vagy kvázi-kezelés) hatására megvalósuló változásáról beszél. Megjegyzem, hogy a filozófiában is található hasonló megközelítés, pl. Patrick Suppes okságfelfogása Hume-ot módosítva a konstans együttjárás kritériumát úgy módosítja, hogy az ok bekövetkeztével az okozat nagyobb valószínűséggel jelenik meg.

Bár Froeyman (2012) a tényellentétes megközelítésbe sorolja, érdemes mégis Wesley Salmon (elsősorban Bertrand Russell-re és Hans Reichenbach-ra támaszkodó) munkáját önállóan tárgyalni itt – elsősorban azért, mert hozzá hasonló található a kauzális következtetések szociológiai típusai között is. E. Szabó László (2004) **ontológiai megközelítésként** definiálja ezt a fizikalista megközelítést, ami az elemi, partikuláris fizikai események közötti kauzális viszonyokra - elsősorban kauzális folyamatokra, s nem egyszerűen kauzális relációkra - koncentrál. A világnak ezt a mélyebb kauzális struktúráját tartja lényegesnek, a komplex eseménytípusok közötti, általa indukált regularitást csak másodlagosnak. Salmon e két szint közötti viszony leírására a „fekete doboz kinyitása”-metaforát használja: „*One can look at the world, and the things in it, as black boxes whose internal workings we cannot directly observe. What we want to do is open the black box and expose its inner mechanisms.*” (Salmon, 1988, pp. 77). Ebben a megközelítésben a kauzalitás ontológiailag tartalmas fogalom: a kauzális kapcsolatot partikuláris fizikai események közötti kölcsönhatások valósítják meg. Salmon egészen explicit az episztemológiai megközelítés felől az ontológiai felé tartó, általa kíváncsnak tartott elmozdulással kapcsolatban: „*To shift from epistemic to the ontic conception involves a radical gestial switch.*” (Salmon, 1988, pp 328). Az alábbi fejezetben látni fogjuk, hogy ez az alapállás a szociológiában megjelent elemzési módszerek közül leginkább a mechanizmusokra hivatkozó megközelítéssel rokonítható.

1.5. Az oksági következtetés négy típusa

1.5.1. Okság, mint robusztus összefüggés, avagy neo-humeiánus regularitás

A rendszeres együttjárás a kiindulópontja a robusztus összefüggésekre hivatkozó tudományos oksági megközelítésnek (Goldthorpe, 2001, terminológiája). Brady (2009) ugyanezt neo-humeiánus regularitás-alapú megközelítésnek nevezi, azaz nem a robusztusságot tesztelni kívánó módszertani aspektust, hanem a Hume-i alapokon nyugvó modern okság-elméletekkel

⁹ Ez persze így egy nagyon elnagyolt kijelentés, a fizikával szemben a szociológiának pl. hivatkozni kell a vizsgálat tárgyainak tudatára, a cselekvők preferenciáira és információira is, de azt hiszem, a társadalmi világ indeterminisztikus jellege minden definíció mellett elfogadható konszenzuálisan.

való kapcsolatot hangsúlyozza. Utóbbiak a regularitás fogalmát Hume-nak az okságra vonatkozó állandó együttjárás (*constant conjunction*) kritériumához kapcsolják. (Hume másik két kritériumát is elfogadja a megközelítés, ezek szerint az ok időben megelőzi az okozatot ill. ok és okozat térben/időben érintkeznek).

Az okságnak a szociológia kvantitatív alkalmazási gyakorlatában ez a leginkább elterjedt megközelítése, és a bevezetőben idézett problémának (*Correlation does not imply causation*) is ez a leggyakrabban, a tankönyvekben is javasolt feloldása. Kiindulópontja az az állítás, hogy a megfigyelt asszociáció valóban nem jelent feltétlenül oksági kapcsolatot, de az okság valamilyen módon mindig asszociációt implikál¹⁰. Tehát a megfigyelt asszociáció mögött bizonyos esetekben okság állhat, a kérdés az, hogy az X és az őt időben követő Y változók között megfigyelt asszociáció vajon mikor és milyen mértékben feleltethető meg az X-Y-ra gyakorolt kauzális hatásának. A probléma megoldására a megközelítés az asszociáció robusztusságának tesztelését ajánlja: akkor feleltethető meg az asszociáció oksági kapcsolatnak, ha újabb változók bevonása után (műszóval ezek „kontrollálása”, *controlling* révén) is megmarad az asszociáció. A kontrollálással számolt asszociáció az ún. parciális asszociáció, ezért nevezik ezt a módszert máshol (pl. Cook and Campbell, 1979, pp. 26) parcializálásnak (*partialling approach*). Azzal kapcsolatban, hogy mely változókra érdemes/kell kontrollálni, a megközelítés általában nem fogalmaz világosan; e probléma explikálása és megoldása Pearl nevéhez fűződik (lásd az irányított körmentes gráfok kapcsán az 0. fejezetben).

Az oksági kapcsolat definíciójának birtokában ugyanakkor továbbra is kérdés, hogy hogyan bizonyítható e kapcsolat megléte. A módszer nyilván potenciálisan végtelen sok változó kontrollálását kívánná meg, ezért bizonyításra nincs lehetőség, csak a kauzális fontosság melletti (a bevont Z-k számának növekedésével egyre erősödő) érvekre. Ezzel kapcsolatban lásd majd a tényellentétes megközelítés nyújtotta tisztázást erre a problémára a 4. fejezetben.

Ellenkező esetben, ha a kontrollálás után az asszociáció eltűnik, az asszociációt hamisnak (*spurious*) tekintjük. A bevett angol szóhasználat: tényleges ok (*genuine cause*) vs. hamis ok (*spurious cause, fictitious cause*). Egy korábbi példám kifejtéseként: ha X az adott ország házassági rátája, Y az országban adott időszakban írt tweetek boldogságszintje, akkor a köztük levő erős korreláció adódhat abból, hogy az ország kulturális tradíciói (Z) hatnak a házassági hajlamra és a szövegek emocionális töltöttségére is. A Z-re kontrollálva (ennek legegyszerűbb módja az, ha azonos kultúrájú országokat tekintek) lehet, hogy a házassági ráta

¹⁰ Ugyanakkor van kivétel, az oksági kapcsolatban levő X és Y közötti asszociációt egy harmadik változó, Z akár ki is olthatja, lásd részletesebben az 0. fejezetet. Ez az eset a Lazarsfeld-től induló elaborációs eljárás egyik (elfedésnek, *suppression*-nek nevezett) alosete: a kétváltozós kereszttábla vizsgálata nem mutat ilyenkor asszociációt, de a megfelelő harmadik változó bevonása felfedi a parciális kapcsolatot. A kioltásra a sokat idézett példa: a fekete és fehér bőrű, gyilkossággal vádolt vádlottak halálbüntetése azonos arányú (nincs kétváltozós kapcsolat), pedig a büntetőeljárás valójában diszkriminatív a feketékkel szemben. Ezt a részletes elemzés fedi fel: mind a fehér, mind a fekete bőrű áldozatok vélelmezett gyilkosaira igaz, hogy a fekete vádlottak kapnak nagyobb arányban halálbüntetést. E diszkrimináció azért kerül elfedésre, mert az esküdtszék az áldozat fehér bőrszínét implicit módon súlyosbító körülményként ítéli meg: fehér áldozatért nagyobb valószínűséggel adnak halálbüntetést, viszont fehéreket inkább a fehérek gyilkolnak meg.

és a boldogság már függetlenek. Az ehhez hasonló szerepben levő, a nyers asszociációt kontrollálás után lényegesen megváltoztató Z változókat *confounder*-ként ismeri a szakirodalom (bevett magyar fordítás nincs rá, néhol összemosó változóként említik). A fogalom jelentésébe legtöbbször impliciten azt is beleértik, hogy a kontrollált asszociáció megegyezik a „valódi hatás”-sal, vagy legalábbis közelebb van hozzá, mint a nyers asszociáció, erre utal a *spurious* és a *fictitious* jelző is.

A fenti példát az 1. ábrán jelölt tudományos megismerés-típusok egyike, az Elster által elkülönített korrelációs magyarázat kapcsán már említettem. Ha a korrelációt nem az eredeti statisztikai jelentésében használom, hanem általánosan asszociációs mértéknek tekintem (Elster ezzel kapcsolatban nem fogalmaz egyértelműen), akkor az okság, mint robusztus összefüggés-megközelítés tudományos magyarázatban alkalmazva megfeleltethető Elster korrelációs elemzésének. A megközelítés azonban nem csak magyarázatban, hanem előrejelzésben is alkalmazható. Granger (1969) idősoros közgazdasági adatokból kiindulva akkor nevezi az Y-t időben megelőző X-et az Y Granger-okának, ha minden egyéb adat birtokában az X értékeinek ismerete pluszinformációt jelent az Y értékeinek előrejelzéséhez.

Fontos megjegyezni, hogy a megközelítés szociológiában megfigyelt elterjedtségének oka az, hogy a társadalomtudományok - a természettudományokkal szemben – tipikusan sajátos empirikus anyagból kell, hogy kiinduljanak. Ez az anyag egyedi események megfigyelésén alapszik, mely események létrejöttére, létrejöttük körülményeire – a kísérletes természettudományokkal szemben – a megfigyelő nem lehet befolyással. Ezt már Durkheim (1894) is expliciten megfogalmazta a dolgozat bevezetőjében idézett tanulmányban. Durkheim gondolata a kísérletek és megfigyeléses vizsgálatok egyik első, klasszikus szembeállítás. E munka után három évvel jelenteti meg Az öngyilkosság-ot, ahol megfigyeléses adatokból kiindulva, a megfigyelt együttjárások alapján alkot oksági magyarázatokat.

A Lazarsfeld-féle elaborációs technikaként ismert kutatási program (pl. Kendall, Lazarsfeld, 1950) is ehhez a megközelítéshez sorolható. Az technika lényege, hogy X és Y kategoriális változók kereszttáblájában megfigyeli az (un. nulladrendű) asszociációt, majd egy harmadik, Z változó bevonásával (a Z rögzített értéke mellett kapott feltételes kereszttáblákon) a parciális asszociációt. X és Y kapcsolatát aszerint értelmezi, hogy a parciális kapcsolat a nulladrendűhöz képest hogyan változott (erősödött, csökkent, vagy eltűnt), illetve hogy a Z változó időben a másik kettőhöz képest hol helyezkedik el (közöttük vagy előttük). A módszer általánosított, folytonos változókra is értelmezett megfelelője tulajdonképpen a regresszió- és az útelemzés (és utóbbi modern megfelelője, a strukturális egyenletek modellezés) is, melyek direkt és a Z-kre kontrollált indirekt hatások megkülönböztetésére képesek.

A kauzalitás, mint robusztus összefüggés a hazai és nemzetközi szociológiaoktatásban is egyezményesen elfogadott alap, elég csak az egyik legelterjedtebb kutatómódszertani jegyzetet, Earl Babbie sok kiadást megélt, magyarra *A társadalomtudományi kutatás gyakorlata* címmel fordított munkáját felütni. A könyv az oksági kapcsolat három kritériumát sorolja fel: a változóknak korrelálniuk kell, az ok meg kell, hogy előzze az okozatot, ill. a

kapcsolatnak „*non-spurious*”-nek kell lennie, vagyis meg kell maradnia további változók bevonása után is.

Bár a megközelítés, mint említettem, igen elterjedt a szociológiai praxisban, expliciten általában mégsem kapcsolják össze a parciális asszociációk elemzését a kauzalitás fogalmával. Illusztrációként megjegyzem, hogy a Szociológiai Szemlében megjelent tanulmányok között keresve több tucat cikkben találjuk meg a „kontrollálva” kifejezést, de ezek többsége nem tesz említést a vizsgált változók közötti kapcsolat oksági jellegéről, még kérdésfeltevés szintjén sem.

Az oksági kapcsolat definíciójának birtokában ugyanis továbbra is kérdés, mikor, mely változók kontrollálása után tekinthető bizonyítottnak e kapcsolat megléte. Egy szélsőséges, de meglehetősen elterjedt, kritikusai által *kitchen sink regression*-nek nevezett gyakorlat szerint a minél több (mintegy a mosogatóba válogatás nélkül beöntött) kontrollváltozó jelenti a biztosítékot a kapcsolat robusztusságára. De a kevésbé szélsőséges gyakorlat is kritikákat vált ki. Goldthorpe (2001) szerint e kritikák kiindulópontja az, hogy a megközelítés kizárólag statisztikai tesztek függvényévé teszi az okság eldöntését, mindenfajta háttértudás, „*subject-matter*” input igénye nélkül. Meg kell jegyezni, hogy ha nem is nagyon egyértelműen, de Lazarsfeld vagy az útelemezés úttörője, Duncan maga is óvatos volt a pusztán statisztikai modellek szociológiai elmélet nélküli használatával kapcsolatban. Pl. Lazarsfeld az elaborációt egy plauzibilis történetvázlat (*story-line*) mellett ajánlotta működtetni (erről Lazarsfeld első felesége, Marie Jahoda ír 1979-es megemlékezésében). Hasonlóan, a robusztus összefüggés Goldthorpe által összefoglalt kritikái szerint az elemzésben szereplő változókat a cselekvés egy általánosított narratívájában kellene elhelyezni. Ez a narratíva egy „kauzálisan adekvát” folyamatot ír le, valamely jól kidolgozott cselekvéseméletre támaszkodva. A robusztusnak talált kapcsolatot ez a narratíva egészíti ki és magyarázza.

1.5.2. Okság, mint a tényellentétes kijelentés teljesülése

A tényellentétes (vagy kontrafaktuális, *contrafactual*) megközelítés a „Ha az ok nem következett volna be, akkor az okozat sem valósult volna meg” kritériumra cseréli le a neo-humiánusok együttjárás-feltételét. Előnye az utóbbival szemben, hogy szinguláris eseményekkel foglalkozhat, nem kell minden egyedi tényállást egy univerzális együttjárási törvény alá vonnia. Éppen emiatt a kismintás problémák, az esettanulmányok is az alkalmazási területei közé tartoznak. Másik előnye, hogy a kauzális hatás *mértékét* is közvetlenül mérhetővé teszi, a két (az ok bekövetkezése és be nem következése által definiált) eset összevetésével. Ahogyan Holland (1986) többször hangsúlyozza: ebben a megközelítésben egy ok csak egy másik „okhoz” képest értelmezhető, nem beszélhetünk általában vett hatásról.

Az egyik legkorábbi példa a tényellentétes állítások tudományos magyarázatban betöltött szerepével kapcsolatban Max Weber 1906-os írása¹¹ Eduard Meyer könyvéről¹². A könyvben

¹¹ Weber, Max (1985 [1906]): Kritische Studien auf dem Gebiet der kulturwissenschaftlichen Logik. In: *Gesammelte Aufsätze zur Wissenschaftslehre*. 6. kiadás, Tübingen: Mohr, pp 267.

Meyer megválaszolhatatlannak és üresnek nevezi a Bismarcknak az Osztrák-Magyar Monarchia megtámadására vonatkozó döntésére vonatkozó mi-lett-volna-ha-nem kérdést. Weber-nek a dolgozat bevezetőjében idézett véleménye ezzel szemben az, hogy a történelem feladata nem az eseménykrónika, hanem éppen az ilyen kérdések megválaszolása; annak az eldöntése, hogy az adott tényezőnek éppen úgy kellett-e történnie ahhoz, hogy az adott következmény létrejöjjön.

David Lewis, a tényellentétes filozófiai elméletének kidolgozója szerint a tényellentétes „ha az ok nem állt volna fenn, akkor az okozat nem következett volna be” kijelentés igazsága a legközelebbi lehetséges világgal ellenőrizhető, amit úgy definiálhatunk, mint a miénkhöz leginkább hasonlót azon világok közül, ahol az ok nem állt fenn. A módszer gyakorlati felhasználásának éppen az a hátránya, hogy a kauzalitás fogalmát ezzel a közvetlenül meg nem figyelhető legközelebbi lehetséges világgal ragadja meg. Annak eldöntése, hogy a lehetséges világ és annak működése hogyan közelíthető meg, ismét (akárcsak a robusztus függőség esetén a megfelelő kontrollváltozók kiválasztása) a kutató háttérismeretén és döntésein múlik.

A legtöbbet hivatkozott társadalomtudományi alkalmazási példák között van a 2000-es amerikai elnökválasztás *butterfly ballot* néven elhíresült floridai fiaskója. Palm Beach-ben olyan szavazólapot használtak, amin az Al Gore-hoz tartozó választási lehetőség nehezen volt azonosítható. Az utólagos számítások szerint ez a szavazólap okozta, hogy kb. 2400 Gore-ra leadott szavazat félrement, és a választást végül George W. Bush nyerte meg. Ezek a számítások a lakóhelyüktől távol levélben szavazók (*absentee voters*) eredményein alapulnak, akik félreérthetetlen szavazólapot kaptak. A számítások arra hivatkoznak, hogy a két (a választások napján ill. a levélben szavazó) populáció hasonló egymáshoz. Vagyis azt a meg nem figyelhető legközelebbi lehetséges világot, ahol a választás napján nem a hibás szavazólapal szavaztak, a levélben szavazók adataival közelíti meg. Más, tényellentétes érvelést alkalmazó szociológiai esettanulmány pl. az, amelyik Argentína egymást váltó politikai rendszereiben elemzi a közigazgatás sajátos jegyeit, s ezek hatását olyan diszfunkciókra, mint a korrupció (Oszlak, 1997). Egy másik példa egy, az USA-t és Európát összehasonlító munkaerőpiaci elemzés (Elwert, 2002). Kérdésfeltevése, hogy a hagyományosan magasabb európai munkanélküliségi ráta mögött állhat-e igazságszolgáltatási eltérés, lehetséges-e, hogy az USA hatalmas és egyre növekvő bebörtönzött populációja rejtett munkanélküliséget fed el. A legközelebbi lehetséges világ ebben a megközelítésben így fogalmazható meg: hogyan alakulnának az amerikai munkanélküliségi mutatók, ha más lenne az USA-beli igazságszolgáltatás büntetési gyakorlata.

A floridai elnökválasztás példájában tulajdonképpen véletlenül tudták pótolni a legközelebbi lehetséges világot a pótszavazókkal. Más esetben kontrollált kísérletekkel adódhat a helyettesítés: ha a kontrollcsoport vélelmezhetően csak a „kezelést” tekintve tér el a „kezelt”

¹² Bár Weber itt a történelemtudományról beszél, szociológiájának alapjait is megtalálhatjuk a munkában. Ezt támasztja alá Erdélyi Ágnes 2011-es tanulmánya is, ahol Webernek ezt a munkáját elemzi, az adekvát okozás weberi fogalmán elindulva, az írást a szociológia számára relevánsabb 1913-as „A megértő szociológia néhány kategóriájáról” c. munkával összekötve.

csoporttól, akkor a legközelebbi lehetséges világ helyettesítőjének tekinthető.¹³ Ez a hasonlóság vagy randomizációval, vagy a körülmények szigorú kontrollálásával (fizikai kísérletnél a hőmérséklet stb. változatlanul tartásával) érhető el. Megfigyeléses vizsgálatoknál a tényellentétes gondolkodás előnye az, hogy a „mi lett volna ha” kérdéssel a kísérletes szempontot helyezi újra előtérbe, ezzel a megfigyeléses adatok korlátaira hívja fel a figyelmet, és e korlátok máskor implicit átlépésének világosan artikulált indoklását teszi szükségessé.

Azzal kapcsolatban, hogy a tényellentétes megközelítés mely oksági következtetés-típusok esetén használatos, elmondható, hogy pl. az elnökválasztási példa esetén magyarázatot, mégpedig statisztikai jellegű kauzális magyarázatot állítottunk fel Bush győzelmére. A statisztikai jelleg abból fakad, hogy a *butterfly-ballot* csak növeli a téves szavazat valószínűségét, de nem determinálja azt. A történelmi esettanulmányok, mint Bismarck döntésének elemzése, szintén utólagos magyarázatok, de nincs bennük probabilisztikus jelleg, mert egyetlen szinguláris esettel foglalkoznak, viszont nem is hivatkoznak törvényekre, tehát tulajdonképpen nem sorolhatók a kauzális magyarázatok közé. Leginkább a megértéshez lenne sorolható az eset, s valóban, Erdélyi Ágnes (2011) is a racionális cselekvő megértésének eljárásához hasonlítja a történész munkáját, a történelmi szereplők lehetséges cselekvésmódjainak és azok várható következményeinek összevetését. A tényellentétes megközelítés ezen kívül hatásvizsgálatokban is használatos, erre jó példa Holland statisztikai okság-modellje, melyet alább külön fejezetben tárgyalok. Holland modelljének továbbgondolásából több statisztikai megoldás született (pl. *propensity score*-alapú párosítás) a szociológia szempontjából fontos megfigyeléses vizsgálatok tényellentétes kereten belül történő kezelésére. Ezekben a vizsgálatokban hatáserősséget szeretnének mérni, de nem tudják randomizálni a kezelt és kísérleti csoportot. A *propensity score*-alapú párosítás hazai alkalmazására példa Hárs és Simon (2013) vagy Kertesi és Kézdi (2012). Hárs és Simon az OFA 2009-es munkahelymegőrző válságprogramjának hatását méri ki (a mi-lett-volna-ha-nem-nyer-támogatást adja a tényellentétest), míg Kertesi és Kézdi a roma és nem roma tanulók iskolai teszteredményei közti különbségben számszerűsíti a rossz életkörülmények hatását.

1.5.3. Okság, mint hatást kiváltó manipuláció

A manipulációs irányzat a vizsgálat tárgyába történő aktív beavatkozást, az ok manipulációját és a manipuláció hatásának megfigyelését szorgalmazza, kísérleteken keresztül alapozva meg az okságot. Ez az aktív beavatkozás a tényellentétes megközelítés esetén, még ha az kísérleten alapul is, nem tartozik hozzá az okság episztemológiájához. Pedig a manipuláció maga az, ami több fontos, a tényellentétes esetén felmerülő probléma (a hatásirány megállapíthatóságának, a látszólagos korrelációnak és az oksági megelőzésnek, a *preemption*-

¹³ Az idézőjelek szerepeltetésének oka, hogy e fogalmakat nagyon általános értelemben kell venni, a kezelés pl. akár a diploma megszerzése is lehet.

nek a problémája, lásd pl. Elster 1983, pp. 35-40.) megoldását kínálja¹⁴. Alapvető logikai különbség, hogy a tényellentétessel szemben a manipulációs megközelítés tipikusan az okból indul ki, és a hatást kívánja megadni. A randomizált kísérlettel alátámasztott, manipulációs megközelítésű elemzés eredménye utólag sosem cáfolódik meg, szemben a tényellentétes megközelítést használó, megfigyeléses vizsgálaton alapuló magyarázattal – utóbbi eredménye a háttértudás növekedtével módosulhat.

Ahogy Brady (2009) írja, filozófusok ritkán, statisztikusok viszont annál inkább egyesítik a két megközelítést. Goldthorpe (2001) is az utóbbiak közé tartozik. Az egyesítés oka az lehet, hogy a két megközelítés a tervezett kísérletek két aspektusát, a beavatkozást (manipuláció) és a kontrollt („ha az ok nem következik be”) jeleníti meg, és mindkét aspektus egyaránt fontos a kauzális reláció szempontjából. A beavatkozás vizsgálata kontrollcsoport nélkül a hamis korreláció lehetőségét hordozza magában. Ilyen eset lenne, ha a nyugdíjba vonulás pszichológiai hatásait vizsgálnánk az azonos korú aktívak követése nélkül, s a nyugdíjba vonulás hatásaként azonosítanánk azt, ami valójában az idősödés hatása. S fordítva, a kezelt/kontroll összevetés manipuláció nélkül nem adna számot a hatás irányáról. Pl. ha azt találjuk, hogy a pszichés betegek társas támogatottsága alacsonyabb, mint az egészségeseké, kérdés marad, hogy melyik tényező az ok, s mi következik be tipikusan előbb, a pszichés probléma vagy az elmagányosodás.

Társadalomtudományokban elsősorban közpolitikai programok értékelésénél merül fel az intervenciók (valós társadalmi környezetben végzett kísérletek) alkalmazása, gyakran használják pl. az USA-ban új iskolai programok értékelésére. Egy jó példa a társadalomtudományi evaluációs módszerek szakértőjének, Thomas Cooknak társaival készített tanulmánya (Cook et al., 2000), ahol 23 középiskola bevonásával randomizált kísérletben, ismételt mérésekkel vizsgálták egy új program hatását. A manipulációs megközelítéshez tartoznak a szociológiai tervezett kísérletek is, mint Milgram szociálpszichológiai kísérlete, vagy a diszkriminációt mérő kísérletek (hazai példák Sik Endre és Simonovits Bori munkaerő- és lakáspiaci diszkriminációt vizsgáló tervezett kísérletei, ahol kontrollcsoportot is használtak). Ide tartozó hazai példa továbbá Örkény és Székelyi (2007) deliberatív közvélemény-kutatása, melynek során azt vizsgálták, hogy az előítéletes gondolkodás hogyan változik más véleményekkel, érvekkel szembesülve - ez előtte-utána típusú mérés volt, kontrollcsoport nélkül.

A közpolitikai hatáselemzésekben kétségkívül nagyon produktív ez az elemzés, ám a szociológiában ritkán tervezhető randomizált kísérletes design, hiszen a vizsgálatok többsége megfigyeléses. A manipulációs modell azonban ilyenkor is fogalmi keretet biztosít az

¹⁴ Brady ugyan említ egy ellenpéldát az oksági megelőzés manipulációs megközelítésben való megjelenésével kapcsolatban, de nem érzem jónak a példát. A példa szerint az arzént régen használták nemibaj kezelésére, bár mint tudjuk, erősen mérgező. Ha tehát nemibetegek random kiválasztott csoportján alkalmazzuk az arzént, azok a szer hatására fognak meghalni, míg a kezeletlen kontrollcsoport a betegségben. Tehát (Brady szerint tévesen) nullának fogjuk mérni az arzént hatását, pedig van hatása, mégpedig negatív. Én viszont azt gondolom, hogy a nemibetegek populációján valóban nincs a szernek hatása, szemben az általános populációval, ahol valóban negatív a hatás. Tehát a kísérlet nem ad téves eredményt.

értékes oksági következtetések levonhatóságához szükséges feltételek megadásához. Például egy ténylegesen megvalósult, nem randomizált munkaerőpiaci intervenció esetén mérlegelendő a randomizálástól való minden potenciális eltérés, így a szelekciós (a kezelt és a kontrollcsoportba történő nem random besorolás okozta) torzítás, a kezelt csoportba való besorolás vagy be nem sorolás (pl. az alanyok motivációira gyakorolt) nem szándékolt torzító hatása, vagy a munkaerőpiacon időközben lezajlott, a programtól független változások hatása. Ezeknek a releváns faktoroknak az explikálása, mérése és kontrollálása szükséges az elemzés során. Az ehhez használható eszközökről bővebben a következő fejezetben, Rubin és Holland manipulációs megközelítést használó statisztikai modellje kapcsán lesz szó. Ugyanakkor, ahogy Goldthorpe (2001) megjegyzi, ezek után a nem-kísérleti jellegű szociológiai vizsgálatokra a manipulációs megközelítésben már nem áll, hogy „sosem halnak meg” (utalva Holland *Old, replicable experiments never die* bonmot-jára), sőt, épp annyira „ideiglenesek”, háttértudásunk függvényei, mint a robusztus függőség megközelítésben kapott eredmények.

Az okság problémájának nem az oksági magyarázat oldaláról, hanem az okok hatásának, e hatás mérésének oldaláról történő megragadása bár a természettudományok számára természetes lehet, a szociológiában nem feltétlenül találkozunk a kutatók elvárásaival. Jól foglalja össze a kísérletes módszerrel szemben felmerülő ellenérzéseket Örkény Antalnak Simonovits Bori diszkrimináció-méréssel foglalkozó disszertációjára adott opponensi véleményének a következő részlete: „Nevezetesen abban látom a problémát, hogy a diszkrimináció mérésére kialakított módszerek, és ezen belül a kontrollált kísérlet talán valóban képes tetten érni a diszkriminatív cselekvést, ugyanakkor csak nagyon korlátozottan tud értékes magyarázatot adni a cselekvés mozgatórugóira, és egyáltalán, a diszkriminatív magatartás társadalmi beágyazottságára és magyarázatára. Úgy is fogalmazhatok, hogy lehetséges egyfajta »hőmérőként« felfogni a módszert, de nem biztos, hogy ezzel eljutunk a »betegség« mélyebb szociológiai megértéséhez”¹⁵. A megértés igénye, amit itt Örkény hiányol, az általam tárgyaltak közül az okság mechanizmus-alapú megközelítéséhez (lásd később) áll legközelebb.

A manipulációs megközelítés szociológiai fogadtatása sok egyéb kritikát is generált (leginkább az ok manipulálható tényezőként való, leszűkítő definíciója miatt, erről bővebben a következő fejezetben). Összefoglalva: a megközelítés eredeti, ideális kísérleti formájában a robusztus függőséghez képest a szociológiában szűkebb területen alkalmazható, ugyanakkor kétségtelen előnye az utóbbi megközelítéssel szemben, hogy fogalmilag pontosabban tisztázott keretet ad.

1.5.4. A Rubin-Holland modell, mint a tényellentétes és a manipulációs modellt egyesítő statisztikai modell

Említettük, hogy a statisztikában gyakran találkozhatunk a tényellentétes és a manipulációs megközelítésnek a kísérleteken alapuló egyesítésével. Egy ilyen egyesítés a kísérletes

¹⁵ Örkény Antal (2012): Opponensi vélemény Simonovics Bori „A diszkrimináció mérése, különös tekintettel a kontrollált kísérlet módszerére” c. doktori értekezéséről.

statisztikai tradícióból kinőtt legismertebb statisztikai kauzalitás-elmélet, a Rubin-Holland modell (a klasszikus hivatkozások Rubin, 1974 és Holland, 1986). A módszer ideális esetben a Fisher által kidolgozott randomizált kísérleti design-t igényeli, de a megközelítés jóval többet nyújt ennél: egy általános fogalmi keretet, amin belül az oksággal kapcsolatos intuíciónk formalizálhatóvá válik. A modell tehát szokatlan módon nem egy módszer, hanem inkább egy keret a módszerek tárgyalásához - a fogadtatás nem is volt teljesen pozitív, lásd pl. a modellt bemutató egyik alapcikk kritikus hozzászólóit (Holland, 1986, és hozzászólások ugyanazon lapszámban), vagy Arjas (2001) ellenkezését a szerinte felesleges fogalmi absztraktumok bevezetésével kapcsolatban. Ugyanakkor nyilvánvaló válik e fogalmi tisztázás és a kezelt/kontroll csoportot kijelölő mechanizmus érvényességét meghatározó feltételek explicitté tételének előnye, ha belátjuk, hogy segítségükkel a kísérleteken túl a megfigyeléses vizsgálatokból jövő oksági elemzések problémáira is kiterjeszthető az elemzés. Az alkalmazott statisztika ezeket a feltételeket használta az utóbbi évtizedekben a megfigyeléses vizsgálatokra javasolt módszerek (instrumentális változók, propensity-score párosítás stb.) létrehozásakor. A Rubin-Holland modell kiemelten fontos tehát az oksági következtetés legújabb statisztikai módszereinek szempontjából, ezért az alábbiakban külön fejezetben részletezem, Holland 1986-os cikke alapján.

Jelölje U (*units*) a vizsgálati populációt, legyenek elemei u -val jelölve. Jelölje S változó azt, hogy egy adott u kezelésben részesült vagy a kontrollcsoportba tartozott: $S(u)=t$ vagy $S(u)=c$. A kezelést itt megint metaforikus értelemben kell venni: kezelt az, aki az általunk vizsgált oknak „ki van téve” (*exposed*), kontroll az, aki nincs kitéve, vagy más oknak van kitéve. Bár az *exposition* és a *treatment* kifejezések aktív beavatkozásra utalnak, Holland a kezelést eredetileg úgy jellemzi, mint ami létrejöhet a kísérletvezető aktív részvételével, de megfigyeléses vizsgálatban tőle függetlenül is. A manipuláció ugyanakkor megjelenik a modellben, mégpedig annál a fontos kitételnél, hogy a populáció minden egyede potenciálisan bármelyik csoportba tartozhat (*potentially exposable*), ezért személyes tulajdonság (Holland kifejezésével attribútum, *attribute*) nem lehet ok. A kitételnek megfelelően a nem hatása pl. nem értelmezhető ebben a modellben, mert a nem attribútum. Egy pályázat hatásvizsgálata viszont beleillik a keretbe. A kitétel szükségességét Holland azzal indokolja, hogy okozás nem jöhet létre manipuláció nélkül („*No causation without manipulation.*”).

A válasz (*response*) az a jellemző, amit a kezeléssel meg kívánunk változtatni. Y -nal jelöljük ezt a változót, $Y(u)$ az u egyén válasza. Az Y értékét kezelés után mérjük. Mivel Y értéke hipotézisünk szerint függ attól, hogy kapott-e kezelést az illető, két változót kell definiálnunk: $Y_t(u)$ a válasz értéke akkor, ha kezeltük az u -t, $Y_c(u)$ pedig ugyanazon (!) u -ra kapott értéke akkor, ha u nem részesült kezelésben. Ezek után a kezelés u -ra gyakorolt hatása az

$$Y_t(u) - Y_c(u)$$

különbséggel definiálható.

Ez a definíció a modell központi eleme. A hatás a populáció egyedein külön-külön van értelmezve, ezért volt szükség a *potentially exposable* feltételre. Vegyük észre, hogy nincs általános hatás értelmezve, és hogy a modell ún. populációs modell, vagyis nincs valószínűségi modell bevezetve, ami lehetővé tenné a hatás, mint valószínűségi változó

vizsgálatát. Továbbá érdemes felfigyelni arra (ezt Holland sokszor hangsúlyozza), hogy itt a t , mint ok hatása csak egy másik okhoz képest (c) értelmezhető, vagyis két ok kell a hatás kiméréséhez.

Van egy fontos feltétele a modell jóldefiniáltságának, amiről nem tesz említést a cikk, pedig azt a kísérlettervezéssel foglalkozó szakemberek (pl. Cox, 1958) már az '50-es években megfogalmazták: ez a kezelés egyedenként stabil értékének feltevése (*Stable Unit Treatment Value Assumption, SUTVA*). A hatás fenti értelmezése ugyanis előfeltételezi, hogy a kezelés adott egyeden kimért hatása független attól, hogy más egyedek megkapták-e a kezelést. Vagyis e feltétel szerint a kezelés hatása egymástól szeparált, interakcióban nem levő egyedeken mérhető. Pedig – különösen a társadalomtudományokban - a hatás erősen függhet attól, hogy az egyed környezetében hányan kapták meg a kezelést, lásd pl. egy új pedagógiai program bevezetésének esetét (az eredményesség függhet attól, hogy az osztály minden tagja részt vesz-e benne), illetve pl. egy innováció elterjedhet nem-kezeltekre is. Rosenbaum (1999) erre a problémára azt ajánlja, hogy az elemzési egységet definiáljuk úgy, hogy azok között már ne legyen inferencia, pl. pedagógiai program esetén ne a tanuló, hanem az osztály legyen az u megfelelője. Más esetekre ez véleményem szerint nem jelent megoldást, pl. a nyereségek végessége esetén: egy pályázaton támogatott cég eredményessége függ attól, hogy versenytársai is támogatást nyertek-e, egy programban valamely hiányszakmára kiképzett munkanélküli elhelyezkedési esélye függhet attól, hány másik pályázó nyert kiképzést ugyanabban a térségben és így tovább. A SUTVA nem-teljesülése úgy is előállhat, hogy a kontrollcsoport - tudomást szerezve a kezelés tartalmáról - megváltoztatja viselkedését. Szociálpolitikai példát mutatnak rá Mauldon és társai (2000): a kaliforniai kísérletben segélyben részesülő kamaszlányok iskolai teljesítményét próbálták javítani, úgy, hogy csökkentették a segélyt, ha a lányok megbuktak az iskolai vizsgákon. A kontrollcsoportból (ismerősi kapcsolatok révén) sokan tudomást szerezhettek a kezeltekre alkalmazott büntetésről, ezért keményebben tanultak, ami által a büntetés hatását a vizsgálat vélhetően alulbecsülte.

A hatás definíciójából látható az is, hogy miért a tényellentétes alá sorolható a modell: ugyanazon u -t tekintjük két különböző állapotban. A modellnek a filozófiai tényellentétessel szemben nincs szüksége a legközelebbi lehetséges világok bevezetésére, mert az oknak kitettséget egyfajta beavatkozásként határozza meg (így nem merül fel az a fontos filozófiai kérdés sem, hogy hogyan épül fel a legközelebbi lehetséges világ, mi az, amiben megegyezik a mi világunkkal stb). Ugyanakkor a modell nyilvánvaló problémája (Holland kifejezésével: *Fundamental Problem of Causal Inference*) éppen ebből ered: lehetetlen **megfigyelni** Y_t és Y_c értékét egy időben ugyanazon az u -n, vagyis lehetetlen **megfigyelni** t hatását u -ra ().

Csoport	Y_t	Y_c
„Kezelt”	Megfigyelhető	Tényellentétes
„Kontroll”	Tényellentétes	Megfigyelhető

2. ábra Az Oksági Következtetés Fundamentális Problémája

Fontos megérteni, hogy e fundametális probléma léte nem a kauzális állítások lehetetlenségét jelenti, csak azoknak az általában **nehezen tesztelhető feltételek bevonásának** (vagyis némi bizonytalanságnak) **az elkerülhetetlenségét**, amikkel a problémát az oksági elemzésekben át szokás hidalni. A Rubin-Holland modell lényege éppen az, hogy ezek a feltételek explicitté válnak, és hangsúly kerül legalább részleges tesztelésük szükségességére.

A fundamentális probléma egyik szokásos tudományos megoldása az egyedek homogenitásának, vagyis a kezelés ill. kezelésből való kimaradás által kiváltott válasz egyedenkénti azonosságának feltételezése:

$$Y_t(u_1) = Y_t(u_2), Y_c(u_1) = Y_c(u_2) \text{ valamely } u_1, u_2\text{-re.}$$

Nyilván ha ez fennáll, akkor a fundamentális probléma eliminálódik, hiszen az u_1 -et u_2 -vel helyettesíthetjük akkor, amikor a kezelésből való kimaradását vizsgáljuk. Az egyedek homogenitása pl. fizikai eszközök esetén valóban plauzibilis feltevés. Bár a hétköznapi gondolkodás sokszor társadalmi példákra is elfogadhatónak tartja (vegyük példának azt a szülőtársamat, aki a Waldorf-módszer előnyét igazolhatónak látta a Waldorf-iskolákba járó gyerekek jobb teljesítményével), a feltevés a szociológiában csak további feltételekkel alkalmazható ténylegesen (az egyedek homogenitását garantáló háttértényezők ismerete kell hozzá).

Egy másik tudományos megoldás a válasz időbeli stabilitásának (mindegy, hogy mikor történik a kezelés és az azt követő mérés) és a hatás átmeneti jellegének (nem befolyásolja az Y_t -t a c -nek való korábbi kitettség, biostatistikai megfogalmazásban: a kezelés hatása idővel „kimosódik”) a feltételezése. Ezen feltételek mellett a fundamentális probléma ismét megkerülhető: ugyanazon egyeden előbb a c -t alkalmazom, lemérem $Y_c(u)$ -t, majd a t -t alkalmazom, és $Y_t(u)$ -t mérem le. Ehhez a megközelítéshez tartozik Örkény és Székelyi (2007) már hivatkozott előtte/utána méréses deliberatív kísérletes módszere is. A természettudományokban ezek a feltételek is könnyebben alátámaszthatóak, de a hétköznapi gondolkodás is gyakran használja őket impliciten (pl. amikor önkormányzati választások kampányában az aktuális polgármester teljesítményét ellenfelei az eltelt négy év bűnözési statisztikáinak romlása alapján ítélik meg). Vagy ide tartozik a csökkentett zsírfogyasztás utóbb tévesnek bizonyult diétájának egyik téves bizonyítéka is: a zsírfogyasztás csökkenésével párhuzamosan az USA-ban valóban visszaesett a szív-érrendszeri halálozás, de feltehetőleg nem a táplálkozás megváltozása miatt, hanem azért, mert orvostudomány is fejlődött. Ugyanis a szívbetegségek száma nem csökkent párhuzamosan (Taubes, 2001).

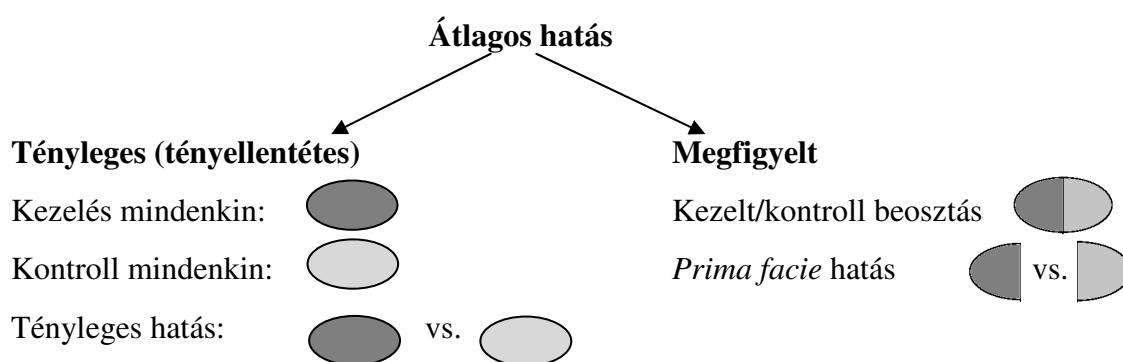
A fundamentális probléma egy harmadik megoldása statisztikai megoldás, s ez felel meg leginkább a szociológiai empirikus gyakorlatnak. Nem az egyénenkénti, hanem csak az átlagos hatást, a

$$T := E(Y_t) - E(Y_c)$$

különbséget kívánjuk ilyenkor mérni.¹⁶ Csakhogy a fundamentális probléma miatt ez nem mérhető, mert a t -re és a c -re adott válasz nem figyelhető meg a populáció összes tagjára, hanem csak a kezelt/kontrollcsoportra, tehát csak

$$T_{PF} = E(Y_t | S=t) - E(Y_c | S=c)$$

mérhető. A T_{PF} itt a *prima facie* átlagos oksági hatás jelölése (ez az okság irodalmában ismert kifejezés Patrick Suppes-től jön, és arra utal, hogy első megközelítésben ezt a hatást szokás tényleges átlagos oksági hatásnak venni). Illusztrációként lásd a 3. ábrát. Az, hogy csak a sötét/világos félkörökben tudjuk mérni a kezelés/kontroll hatását, pl. az egyedek fent említett homogenitása esetén nem jelentene gondot: ilyenkor az egyedek szabadon helyettesíthetők másokkal.



3. ábra A tényleges és a *prima facie* átlagos hatás

A *prima facie* átlagos oksági hatás tulajdonképpen azonosítható azzal az első lépésben mért asszociációval is, amit a robusztus összefüggés-megközelítés az analízis további lépéseiben újabb változók bevonásával kontrollálni igyekszik. T és T_{PF} általában nem egyenlő (a kezelt csoportban tapasztalt válasz nem feltétlenül ugyanaz, mint amit a teljes populációban kapnánk), ám a kezelt/kontroll kiosztás és az Y válaszok függetlensége mellett egyenlővé válnak. Vagyis \perp -nal jelölve a függetlenséget, és $|$ -nal a feltételt, ha

$$(Y_t - Y_c) \perp S \quad (1)$$

akkor

$$E(Y_t | S=t) = E(Y_t)$$

és

$$E(Y_c | S=c) = E(Y_c)$$

amiből

¹⁶ Ez az egyszerű matematikai meghatározás bizonyos statisztikai módszerek esetén komplexebb képet mutat, pl. Rosenbaum (1984) foglalkozott a Blau-Duncan modell és a Rubin-modell összeegyeztetésével, úgy, hogy bizonyos lineáris modellekben az útegyütthatókat megfeleltette az átlagos hatásnak.

$$T_{PF}=T.$$

A 3. ábra jelöléseivel: a sötét és a világos félkörben mért átlagos hatás ilyenkor megegyezik a teljes körökben mért átlagos hatásokkal, joggal helyettesítjük őket velük.

Itt tárgyalandó a szociológia szempontjából fontos nem-randomizált, megfigyeléses vizsgálatok esete. Ezekben a vizsgálatokban szintén átlagos hatásérősséget szeretnének mérni, de nem tudnak randomizálni. Implicit módon a függetlenség feltevését alkalmazzák, a kezelést megelőző változókat használva arra, hogy ezt a feltevést feltételes függetlenségi feltevéssel helyettesítsék. Ilyen feltevés az erős ignorálhatóság (*strong ignorability*) feltevése, ami az okság modern statisztikai megközelítései (pl. propensity score párosítás) használnak. A feltételes függetlenség itt néhány olyan háttérváltozó (X) azonosítását igényli, amikre feltéve a kezelt/kontroll kiosztás (S) már független lesz a potenciális választól (Y_t -től és Y_c -től):

$$(Y_t, Y_c) \perp S \mid X. \quad (2)$$

(Az *erős* ignorálhatóság az Y válaszpár eloszlásának feltételes függetlenségét jelenti a kiosztástól. A *gyenge* ignorálhatóság csak az Y válaszok egyenkénti feltételes függetlenségét kívánná meg.)

Pl. egy munkanélküliek újrafoglalkoztatását célzó, önkéntes jelentkezésen alapuló program hatásvizsgálatakor nem tekinthetünk el attól, hogy a programra jelentkezők valószínűleg lényeges jegyekben (motiváltság, informáltság) különböznek társaiktól, és ezek a jegyek munkakeresésük sikerességét is befolyásolják, ami azért jelent gondot, mert így a programra történő jelentkezés önmagában, a program tartalmától függetlenül is összefügg a munkakeresés sikerességével, tehát felül fogjuk mérni a program hatását. Ám ha azonosítani és operacionalizálni tudjuk ezeket a jegyeket (jel. X változóhalmaz), akkor a rájuk történő kontrollálás után már feltehető, hogy a programra történő jelentkezés nem függ a munkakeresés jövőbeli sikerességétől. Az erős ignorálhatóság feltevésével a kezelt és a kontroll csoportban mért válaszok különbsége (T_{PF}) a tényleges átlagos hatással, T -vel lesz egyenlő.

Az X -re történő feltételes elemzés nehézségekbe ütközik, ha X több változóból áll. Rubin (1997) mutatta meg, hogy X helyett elég az X függvényeként előálló *propensity score*-ra mint feltételre végezni az elemzést:

$$(Y_t, Y_c) \perp S \mid P(D=1 \mid X),$$

ahol $P(D=1 \mid X)$ a kezelt csoportba történő allokálás X mint feltétel mellett előálló valószínűsége, a propensity score. A propensity score ismeretlen, a megfigyelt mintából becsülhető. A kezelt és megfigyelt csoport X szerinti eltérése tehát a propensity score segítségével egyensúlyozható ki. A párosító (*matching*) módszerek ezt a logikát követve úgy kontrollálnak az X -re, hogy az azonos propensity score értékeket felvevő kezelt (u_1) és kontroll (u_2) egyedeket párosítják, és a rájuk számolt hatást, $Y_t(u_1)$ - $Y_c(u_2)$ -t átlagolják. A propensity score ugyanakkor használható többváltozós regressziós modellekben kontroll-változóként, vagy rétegeket képezve belőle rétegzett elemzést is végezhetünk.

Az (1) és (2) egyenletek összevetéséből jól látszik az X tényezők szerepének fontossága. Nyilván az ignorálhatóság feltételének teljesülése sem igazolható az adatokra támaszkodva. A feltétel teljesülésének elfogadhatósága a megfigyelt X informativitásán, gazdagságán múlik. Így akárcsak a robusztus összefüggés kontrollváltozói esetén, ismét a háttértudás szerepéhez jutunk: az ignorálhatóság statisztikai módszerekkel nem igazolható. Ugyanakkor közvetett módon tudunk statisztikai módszerekkel dönteni az elfogadhatóságáról: érzékenységtesztet végezhetünk a konkrét kutatási kereten belül arra vonatkozóan, mennyire befolyásolná a kapott eredményeket egy nem mért, látens *confounder* (Rosenbaum, Rubin, 1983, Cornfield, 1959, és egy izgalmas orvosi alkalmazás: Connors et al., 1996).

Az átlagos hatásérősség sok vizsgálati probléma esetén megfelelő mérőeszköz. Kérdés persze, hogy mekkora a hatás szórása. Ha a szórást nullának tételezzük fel, akkor a konstans hatás feltevéshez jutunk, ilyenkor minden u egyedre megegyezik az $Y_t(u)$ - $Y_c(u)$ értéke. Ez pl. az egyedek homogenitásakor is teljesül, ezért a konstans hatás feltevés a homogenitás-feltétel gyengébb változatának tekinthető. Holland megállapítása szerint Hume okság-felfogása beleilleszthető a modellbe a konstans hatás feltétel hozzáadásával (Hume másik két feltétele definíció szerint teljesül a modellben). Holland tehát nem állítja szembe saját okság-felfogását a humeianus vagy neohumeianus felfogással, csak megadja azt a formális bázist, amin az eltérésük értelmezhető.

Összefoglalva: a modell egy egyszerű definíciót kínál a hatásra, de a fundamentális probléma miatt a hatás mérése csak további feltételek bevonása révén lehetséges, mégpedig vagy 1) a SUTVA és a) az egyedek homogenitása vagy b) időbeli stabilitás és kimosódás ill. 2) a SUTVA és a függetlenség/feltételes függetlenség feltevése révén. A SUTVA el nem hagyható feltétel; nem teljesülése esetén lehetetlen a kísérleti eredményt általánosítani, de sok esetben még értelmezni is. A SUTVA nem-teljesülése esetén a legközelebbi lehetséges világ megkonstruálása nem lehetséges, a randomizációhoz hasonló módszerekkel sem. Nyilván nincs direkt teszt, ami a SUTVA teljesülését vizsgálná; a feltétel teljesülésének támogatása a kísérlettervezés során, gyakorlati eszközökkel lehetséges: az egyedek elszigetelésével, a kezelések és egyéb körülmények uniformizálásával.

A Rubin-Holland modellt a gyakorlatban is alkalmazhatóvá tevő feltételek tárgyalása után a modell szociológiai alkalmazhatóságára, ill. annak legfontosabb akadályára, a modell legtöbbet (pl. Glymour ill. Granger a Holland 1986-os cikkére íródott hozzászólásában) támadott kitételére, a lehetséges okok leszűkítésére térnek ki. Láttuk, hogy a modell az okot potenciálisan kezelésként alkalmazható faktorként definiálja. Attribútum (meg nem változtatható tulajdonság) nem lehet ok, véli Rubin és Holland. *No causation without manipulation*.

A következő példák jól mutatják a Rubin-modell ezzel kapcsolatos, véleményem szerint problematikus korlátait. Holland szerint a "Jól vizsgázott, mert lány" kijelentésben **a női mivolt attribútum**, ezért a kijelentés nem a modellben definiált kauzalitásra vonatkozik, csak egy, a populációban megfigyelt együttjárásra¹⁷. Holland szerint hasonló okokból pl. az iskolai

¹⁷ Gerő Mártonnak köszönhetem a következő, a fentieknek némiképp ellentmondó példát: „Megölte a csecsemőt, mert lány.” Itt az az érzésünk, hogy a kijelentés világosabban értelmezhető kauzálisan. Az anya szemszögéből,

teljesítmény sem lehet ok, ezzel a mobilitáskutatás kérdéseinek jó részét kauzalitás szempontjából szintén értelmezhetetlenné nyilvánítja. Bár mindez a szociológiailag értelmes oksági kérdésfeltevések szükségtelen szűkítésének tűnhet, akad, aki egyetért vele: Zuberi (2001) a faj (*race*), mint attribútum okként való szerepeltetésének elkerülését javasolja. Szerinte a faj és más tényezők között inkább együttjárásról lehet beszélni, mely az egyébként triviálisan meglevő faji alapú rétegződést írja le. A valódi kérdés szerintem nem ez az együttjárás, hanem a faj társadalmi megítélése, ennek történelmi és társadalmi kontextusa. Az együttjáráson túli kauzális kapcsolat feltételezése maga is a faji alapú rétegződés igazolására szolgálhat - véli.

“Jól vizsgázott, mert sokat tanult” – ez a mondat is problémás Holland szerint, mert ugyan a **tanulás nem attribútum, de önkéntes dolog**, nem készíthetünk valakit rá. Vegyük észre, hogy itt a magyarázat a kauzálisból a bevezetőben említett intencionális felé csúszik, a „Jól vizsgázott, mert sokat tanult” és a „Sokat tanult, hogy jól vizsgázzon” mondatok egymás megfelelői a cselekvésnek a cselekvők által tulajdonított értelem szempontjából.

A szociológiai megfigyeléses vizsgálatokban általában sem feledkezhetünk meg a vizsgálati alanyok saját törekvéseiről: tisztában vannak a kísérletben való részvétellel, tudják, mi annak a célja, ezt összevetik saját érdekeikkel és ennek megfelelően cselekszenek. Például a romák szegregált oktatását ellenző középosztálybeli szülők a körzetesített oktatási rendszer integrációs célját átlátva új stratégiákat dolgozhatnak ki (mint ahogyan erre több példa akad Magyarországon) a szegregáció fenntartása érdekében.

A kérdés, miszerint a dohányzás rizikófaktor-e a tüdőráknak (ez a kérdés és a rá adott ellentmondásos válaszok alapozták meg a biostatistikát R. A. Fisher, R. Doll és kortársaik által az '50-es, '60-as években) szintén a modellen kívül esik, hiszen lehetetlen/**etikátlan lenne** a vizsgálati alanyokat cigarettázásra készíteni, ill. megtiltani nekik a dohányzást. A nem-attribútum jellegű, elvben lehetséges, de a **gyakorlatban kivitelezhetetlen** kezelések fontos szociológiai problémáknak a modellben történő tárgyalását is megakadályozzák, pl. Blossfeld és Mills (2001) kérdésfeltevését arról, van-e hatása élettársak esetén a nő teherbe esésének a pár összeházasodására. A terhesség nyilván nem lehet *treatment*, nem osztható ki és nem vonható meg senkitől.

Holland kijelentésének harmadik változata, a “Jól vizsgázott, mert a tanára külön gyakorolt vele” viszont már szépen illeszthető a modellbe, hiszen a tanárral végzett gyakorlásban bárki részesülhet vagy bárkit kizárhatunk belőle.

Hogy a probléma súlyát ne becsüljük felül, fontos megjegyezni: az, hogy a szociológiában vizsgált okok zöme attribútum, ezért a Rubin-Holland modellben nem tekinthetők oknak, az okság fogalmát érinti csak. Ez nem jelenti azt, hogy a modell változatai pl. a feltételes

motivációiból magyarázzuk a viselkedést (szemben a vizsgázással, ahol egy objektív magyarázatot próbálunk adni, s a vizsgázó szempontja is előtérben van). Így talán a lány mivolt könnyebben manipulálható: az a kérdés, mi történt volna, ha kicseréljük a lánygyermeket egy fiúra. Mintha a csecsemő egyetlen lényeges attribútuma itt a neme lenne, ezért nem az attribútum, hanem maga a csecsemő lenne érvényesen lecserélhető.

függetlenségi feltételek explikálásával nem lennének jól használhatók e szociológiai kutatási kérdések megválaszolásában.

Vegyük észre, hogy a szociológiai kérdésfeltevések fent említett problémáinak mindegyike a természettudományos megközelítés társadalomtudományira való átfordításának nehézségéből ered. Goldthorpe (2001) két eltérő választ említ, melyeket a szociológia adott a problémára. Sobel (1995) szerint ragaszkodjunk a Rubin-Holland modellhez, és próbáljuk meg kísérleti vagy kvázi-kísérleti kontextusban megfogalmazni kérdéseinket. Lieberman (1985) viszont úgy gondolja, hogy ez a szcientizmus - a kauzalitás pusztán formális, nem tartalmi megragadása - elfogadhatatlan. A szociológiának saját megközelítést kell találnia, problémáihoz és realiztikusan kivitelezhető vizsgálataihoz illőt. A következőkben ismertetendő megközelítés egy ilyen sajátos szociológiai megközelítést mutat be.

1.5.5. *Okság, mint mögöttes mechanizmus, vagy generatív folyamat*

Az okság itt hivatkozott alapját Goldthorpe (2001) generatív folyamatnak, Brady (2009) mechanizmusnak nevezi. Goldthorpe ezt az oksági megközelítést tartja leginkább elfogadhatónak. Alapállását többen (pl. Collier et al, 2004) úgy interpretálják, mint a kvalitatív argumentum bevonását a statisztikai megközelítésbe. Goldthorpe ebben nem explicit, de valóban, pl. az alábbi idézet arra utal, hogy a csupasz statisztikai kritériumok véleménye szerint kvalitatív jellegűekkel bővítendő: „Az okság ezen felfogásának kidolgozásához statisztikusok is sokféleképpen hozzájárultak. Ugyanakkor ez az irányzat az okság más felfogásaihoz képest kevésbé tükröz statisztikai gondolkodást. Ellenkezőleg: úgy tűnik, éppen annak megragadására tett kísérlet, hogy *mit kell bármely statisztikai kritériumhoz hozzáadni ahhoz, hogy az okság meggyőző bizonyítékát kapjuk.*” (Kiemelés tölem.) Vagyis nem a korábbiakkal szembenálló, hanem azt egy újabb aspektussal való kiegészítésként értékeli a generatív folyamatra történő hivatkozást. Érdeemes megjegyezni, hogy a statisztikusok egy jó része (pl. David Freedman¹⁸) maga is gyakran hangsúlyozta ennek az aspektusnak a fontosságát. Örkény (2012) már idézett kritikája a szociológiai kísérletek korlátozott lehetőségeiről (csak lázmérők, de a betegséget nem találják meg) is a kvalitatív jellegű magyarázatot hiányolja az oksági hatás mérésén túl. Fogalmazhatunk úgy, hogy a manipulációs megközelítés (de tulajdonképpen a robusztus függőség és a tényellentétes is) az **oksági hatás** próbálja megfogni, míg a mechanizmus-alapú megközelítés egy olyan **oksági elmélet** felállítására törekszik, ami az oksági hatás mögött álló társadalmi folyamatokat írja le. Az oksági elmélet az oksági hatásról megszerzett empirikus információkra támaszkodik, és előrejelzések is tehetők segítségével.

Tehát az előbbi megközelítések kiegészítőjeként tekinthetjük a mechanizmusokra hivatkozó oksági fogalmat¹⁹, ugyanakkor azok kritikus pontjait is korrigálhatjuk vele. Goldthorpe pl.

¹⁸ Freedman (2010) a pusztán statisztikai modellek helyett a “shoe leather” oksági vizsgálatot, vagyis a valósággal minél közvetlenebb kapcsolatot fenntartó, a vizsgálat tárgyát a cipő elkoptatása árán is megismerni igyekvő kutatást ajánlja.

¹⁹ Fontos megjegyezni, hogy bár Brady és Goldthorpe a mechanizmusok szerepét nagyon határozottan a szociológiai kauzális magyarázatban látja, a mechanizmusokat valóban alkalmazó irányzat képviselői nem

említi, hogy ebben a szemléletben nincs probléma a „Jól vizsgázott, mert tanult” kijelentéssel, sőt, ez itt standard érvelésnek tekinthető. Valóban, a cselekvők intenciói, melyek a manipulációs megközelítésnek problémát okoztak, a társadalmi cselekvésekre épülő mechanizmusoknak éppen az alkotóelemét adják.

Itt ismét, akár csak a robusztus függőségnél, a hatások okának keresésén (és nem okok hatásának mérésén) van a hangsúly, tehát a kauzális magyarázathoz, azon belül is a deduktív-statisztikai magyarázathoz sorolható a megközelítés. Elster (1983) vélhetően éppen e megközelítés pozicionálása céljából különítette el a korrelációs elemzést az alacsonyabb szintre történő visszavezetést feltételező deduktív-statisztikai magyarázattól. Hedström (2005) szintén a deduktív-nomologikus és az induktív-probabilisztikus magyarázat melletti harmadik magyarázat-típusként határozza meg a mechanizmusokra hivatkozó elemzést. Ugyanakkor a leggyakrabban alkalmazott módszer, az ágens alapú szimuláció nemcsak magyarázatra, hanem előrejelzésre is használható a mikromechamizmusokat leíró paraméterek megváltoztatásával - lásd pl. Elsenbroich összefoglalóját (2012).

A mechanizmusokra hivatkozó szociológiai megközelítés oksági reláció detektálásakor nem elégszik meg azzal, hogy ok és okozat aszimmetriáját az időbeli megelőzéssel biztosítsa (mint a robusztus függőség), hanem a hatás mögött álló mechanizmusokat (más szerzőknél: generatív folyamatot²⁰) igyekszik megtalálni, egy alacsonyabb elemzési szintre történő visszavezetéssel. Ez az alacsonyabb, a társadalmi cselekvések szintjére (individuumokra, relációikra és cselekvéseikre) hivatkozó magyarázat a módszertani individualizmus keretébe illeszkedik. Mechanizmus lehet pl. az önbeteljesítő jóslat vagy a kognitív disszonanciacsökkentés.

A mechanizmusokra épülő irányzat képviselői (pl. Hedström és Ylikoski, 2010) gyakran hivatkoznak a csupán statisztikai mutatókra épülő magyarázatok „fekete dobozának” megnyitására. Ugyanezt a metaforát használja az okság filozófiai megközelítésének Wesley Salmon-féle (elsősorban Betrand Russel-re és Hans Reichenbach-ra támaszkodó) irányzata, melyet E. Szabó László (2004) **ontológiai megközelítésként** definiál. Ez a fizikalista megközelítés az elemi, partikuláris fizikai események közötti kauzális viszonyokra - elsősorban kauzális folyamatokra, s nem egyszerűen kauzális relációkra - koncentrál. Ebben a megközelítésben a kauzalitás ontológiailag tartalmas fogalom: a kauzális kapcsolatot partikuláris fizikai események közötti kölcsönhatások valósítják meg. A fekete doboz-metaforán túl tehát más szempontból is rokonítható e két (szociológiai ill. filozófiai)

hivatkoznak feltétlenül kauzalitásra. Pl. Hedström szerint a kauzalitást nem határozhatjuk meg mechanizmusokkal, mert az utóbbiak definíciója az előbbit is tartalmazza. Ehelyett a tényellentétet javasolja az okság megközelítésére. A megközelítés egyik analitikus eszköze, az ágens alapú szimuláció kauzális magyarázatokra való alkalmasságáról is vita bontakozott ki az utóbbi évtizedben, jó összefoglalóként lásd Elsenbroich (2012) pro-oldali hozzászólását.

²⁰ A statisztikus David Cox vagy Nanny Wermuth *data-generating process*ként hivatkozik ugyanerre, David Freedman mechanizmusként. Brady tipológiájában mechanizmusokat és *capacities*-t említ. Utóbbiak a dolgok/események olyan kauzális képességei, melyekkel állapotváltozásokat, események bekövetkezését képesek előidézni.

megközelítés. Mindkét megközelítés a formális (előbbinél statisztikai, utóbbinál episztemiológiai) tárgyalás helyett tartalmi elemzést kínál, elemi entitásokra bontva a magasabb vizsgálati egységet, az elemi entitások között nem kauzális relációkra, hanem kauzális folyamatokra koncentrálva, az általuk indukált, magasabb szinten megfigyelhető együttjárásokat másodlagosnak tartva.

A mechanizmusokra hivatkozó megközelítésnek több irányzata is elkülöníthető a szociológiában, ilyen pl. a társadalmi kapcsolatháló-elemzés, a racionális döntéshozatal, a játékelmélet vagy az analitikus szociológia (Szántó, Takács, 2012). Az analitikus szociológia meghatározó jegyei világossá teszik, hogyan működhet pontosan egy mechanizmus-alapú megközelítés (Hedström és Ylikoski, 2010): a) a magyarázandó jelenség mindig valamely kollektív társadalmi jelenség, b) a magyarázat cselekvéshozatalra hivatkozik, c) mechanizmusokon, mint középszintű elméleteken alapulva, d) ágens-alapú szimulációt alkalmazva. Az utóbbi modellezés olyan szimulációt jelent, amely a magyarázandó komplex jelenség részeit, a cselekvőknek a modellben feltett relációin, tulajdonságain és feltételezett cselekvési döntésein alapszik. A hagyományos, hipotézisvizsgálaton alapuló statisztikai eljárásokkal szemben ez a módszer jól illeszthető a mechanizmusok vizsgálatához, magára a szimulációra pedig azért van szükség, mert a vizsgált jelenségek komplexitása nem tenné lehetővé a direkt megközelítést. A szimulációk nem tekinthetők valódi bizonyítéknak, csupán a modell mögött álló feltételek következményeinek megtalálására szolgálnak.

Schelling híres szegregációs modellje (1978) jól illusztrálja az ágens-alapú modellezést. Egy 8x8-as táblára, mint házakba két eltérő bőrszínnel rendelkező lakókat helyezett, ahol adott lakót költözésre készítetett az, ha egy (előre rögzített, mindenkire érvényes) toleranciaküszöböt elért szomszédjai között az ellenkező bőrszínűek aránya. Schellingnek ezzel a modellezéssel sikerült megmutatnia, hogy a szegregáció felléphet akaratlan következményként is (még ha az egyéneknek nincs is erős preferenciája a kevert etnikumú szomszédság elkerülésére, hiszen a 60% meglehetősen magas toleranciaszint); s azt is, hogy a mikroszinten bekövetkező apró változások (a városmodell fenti jegyeinek változtatása) hatalmas különbségeket okozhatnak a makroszinten. Egy másik, frissebb, de szintén sokat idézett példa Bearman és társai kutatása (2004) az amerikai szerelmi és szexuális kapcsolati hálózatokról. A mikro-cselekvés paramétereit addig variálják, amíg a szimulált hálózatok jegyeikben leginkább hasonlóakká válnak a megfigyelt hálózatokhoz (e hasonlóságot statisztikai hipotézistesztekkel alátámasztva). Az így megtalált mikro-cselekvési sajátosságokat azután a járványterjedések megelőzésére tekintettel az egészségpolitika figyelmébe ajánlják.

Itthon pl. Bálint és Bozsonyi (2012, választói magatartás modellezése környezet véleménypolarizáltságának függvényében) vagy Tóth, Hajnáczy és Bozsonyi (2013, Tiszabő roma-szegregációjának rekonstrukciója) használták a módszert.

A példák alapján érthető, hogy a mechanizmusokra történő hivatkozás nem nyújt örökérvényű magyarázatot, például újabb magyarázat merülhet fel egy még alacsonyabb szinten. Viszont a magyarázatok megerősíthetőek vagy falszifikálhatóak, a statisztika e megerősítésben vagy cáfolatban ismét szerepet kap. Azonban véleményem szerint óvatossá kell lenni azzal kapcsolatban, hogy a szimuláció eredményeként a magyarázat mely eleme nyer megerősítést

vagy cáfolatot. Schelling szimulációjával pl. megerősítést nyert, hogy viszonylag liberális toleranciaszint mellett is kialakulhat szegregáció. Ugyanakkor a szimuláció nem teszteli a mikrocelekvés modellben alkalmazott módját. Schelling példájánál maradva: ha ezek a modell paraméterbeállításai, akkor szegregált városhoz jutunk, de ez nem bizonyítja sem az elköltözési modell helyességét, sem azt, hogy ez lenne az egyetlen, szegregációhoz vezető cselekvésmodell. Ugyanez áll a szerelmi kapcsolathálózatok vizsgálata során azonosított mikromechanizmusokra. A gyakran szintén mechanizmusokra hivatkozó hálózattudományból is vehetünk példát. Barabási-Albert László ismert tétele szerint ha egy hálózat önszerveződését a Máté-effektussal, mint mechanizmussal irányítjuk a szimulációban, akkor a létrejövő hálózat – sok valós hálózathoz hasonlóan - skálafüggetlen lesz. Ez nem bizonyítéka a Máté-effektusnak, viszont megerősíti azt, hogy egy egyszerű és ráadásul plauzibilis feltétellel skálafüggetlen hálózathoz juthatunk.²¹

Éppen emiatt fontos az a Kovács és Takács (2003) által is hangsúlyozott kitétel, miszerint a mikrocelekvés modellje nem lehet ad-hoc, hanem elméletre kell, hogy épüljön. Az ágens-alapú modellezés, mint tudományos bizonyítási eljárás szokatlanságát az adja, hogy - Szántó és Takács (2012) szavaival – egyfajta tervezett kísérlet: kísérlet, de nem valós kísérleti alanyokkal. Egyszerre mutat (a kiinduló mikrocelekvési premisszák miatt) deduktív és (a szimulációva nyert, ám nem valós adatok miatt) induktív jegyeket, ezért egyszerre nyújt lehetőséget a komplex rendszerek tulajdonságainak levezetésére és váratlan, rejtett jelenségek felfedezésére (Kovács, Takács, 2003).

Végül egy megjegyzés: szintén a mechanizmusra, generatív folyamatra hivatkozó megközelítéshez sorolható a strukturális egyenletek módszere (SEM) is. A módszer a Wright, Duncan és Blalock-féle útmodell általánosítása. Nyilván inkább statisztikai modell, mint elméleti paradigma, és a szociológián belül még nem igazán elterjedt, de érdemes itt megemlíteni. A módszer olyannyira fémjelzi a mechanizmusokon alapuló paradigmát, hogy Judea Pearl 2000-ben megjelent, az okságot mesterséges intelligencia-kutatás oldaláról tárgyaló alapművében úgy fogalmaz, hogy az okságnak kétféle tudományos megközelítése létezik: a tényellentétes és a SEM. A SEM expliciten modellezi azokat a mechanizmusokat, amelyek az okok és hatások láncát összekötik, és, ahogyan Pearl fogalmaz, a Mi-lenne-ha kérdés helyett a Miért? és Hogyan? kérdésre keres választ. A SEM modell egy, a változók közötti kapcsolatot reprezentáló gráf, ahol a gráf csúcsai között húzódó élek oksági kapcsolatokat jelenítenek meg. Ez a reprezentáció alkalmas az oksági utak, közvetett és közvetlen kapcsolatok megjelenítésére, *confounder*-ek megtalálására és hatásuk kiszűrésére stb. Bizonyos szempontból a módszer robusztus összefüggés-megközelítés (ahol csak a potenciális *confounder*-ek hatásának kiszűrése történik meg, mint pl. a lineáris regresszióban) egy általánosított, kifinomultabb változata.

²¹ Az itt használt módszer, az ágens alapú szimuláció evolúciós biológiában alkalmazott változatának hasonló értelmezési problémáira hívja fel a figyelmet Huneman (2011): ha a madarak idegrendszerének három speciális sajátosságát feltéve a szimuláció a valóshoz hasonló repülési mintázatot ad, írja, az nem e három sajátosságra nyújt bizonyítékot, hanem csak arra, hogy nem szükséges központi irányítást feltételezni az összehangolt repülés létrejöttéhez.

1.5.6. A négy megközelítés összevetése

Mint láttuk, a négy megközelítés eltérő empirikus bázis birtokában feltett szociológiai kérdések megválaszolásában nyújt eltérő tudományfilozófiai alapon segítséget. Elkülönítésük jól megfogható néhány összevetési szempont kiemelésével. Az egyik ilyen szempont **ok és okozat aszimmetriájának** kezelése. A neo-humeiánus megközelítés ezt az időbeli megelőzés feltételével oldja meg, a tényellentétes megközelítés logikai állítások aszimmetriájával (az ok nem feltétlenül elégséges, de szükséges feltétel), a manipulációs megközelítésben az ok manipulálásának megfigyelésével, végül a mechanizmusokra hivatkozó megközelítés az ok és okozat közötti oksági reláció helyett a mögöttes mechanizmus fókuszba helyezésével. Egy másik összevetési szempont **az okból vagy okozatból való kiindulás** kérdése. A neo-humeiánus és a mechanizmusokra épülő megközelítés a hatások okát keresi, míg a manipulációs egyértelműen okok hatására koncentrál. A statisztikai tényellentétes megközelítés, kísérletekre hivatkozva szintén ezt teszi, de hatások okát kereső tényellentétes magyarázatokat is találhatunk, ha a példaként említett komparatív esettanulmányokra gondolunk (a filozófiai tényellentétes is hatások okát keresi). Mindezzel összhangban **tipikus empirikus bázisaik** is különböznek: a neo-humeiánus megközelítés tipikusan megfigyeléses vizsgálatokból indul ki, a tényellentétes kísérletekből és komparatív esettanulmányokból, a manipulációs kísérletekből, természetes kísérletekből és kvázi-kísérletekből, végül a mechanizmusokra épülő analitikus modellekből. Melyik a legmegfelelőbb empirikus bázis? Az arany standard a kontrollált kísérlet lenne, ám a szociológiában (a fent tárgyalt elvi vagy praktikus okokból) ez legtöbbször kivitelezhetetlen. A Rubin-Holland modell segítségével viszont megfigyeléses vizsgálat is felhasználható a tényellentétes vagy a manipulatív okság megközelítésére.

Láttuk, hogy aszerint is különbséget tehetünk a négy megközelítés között, hogy „örök” vagy „ideiglenes” eredményt adnak, ez leginkább azon múlik, hogy szükséges-e tartalmi háttértudás az oksági következtetés felállításához. Holland (1986, viszontválasz) szerint minden magyarázat az utóbbi kategóriába tartozik, és csak a kísérletesen igazolt hatás-mérések adnak örök, háttértudástól független eredményt.

Hume, John L. Mackie és más filozófusok azt vizsgálták, hogy egy tényező lehet-e ok vagy sem. A manipulációs megközelítéshez tartozó hatásvizsgálat, vagy a neo-humeiánus korrelációs elemzés viszont ki is méri az ok hatását. A hatás mértékére vonatkozóan általában statisztikai hipotézistesztek szignifikanciáját hivatkozzák, amivel kapcsolatban viszont felmerül a kérdés: társadalmilag, gazdaságilag is jelentős hatás áll-e a statisztikailag szignifikánsnak bizonyult (gyakran közvetlenül nem is értelmezhető skálán mért) különbség mögött? A szociológiai praxisban e kérdés ritkán kerül elő, s ez párhuzamba állítható a kauzalitás fogalmának már említett óvatos kerülésével: a szakmai szignifikancia a tartalmi oldalt érinti, szakmai háttértudás lenne szükséges a megítéléséhez, mérlegelése nem redukálható statisztikai kérdéssé.

1.6. Szociológiatörténeti vonatkozások

Érdemes lenne áttekinteni, hogy a szociológiatörténet paradigmaticusan változó tartalmi problémáival együtt változott-e a kauzalitás-fogalom is, hogy az alkalmazott módszertan

változása befolyásolta-e ezt a fogalmat, hogy az eltérő megközelítések konfliktusainak melyek voltak a forró pontjai, vagy hogy hogyan hatottak a más tudományok felől, kívülről jövő hatások a kauzalitás megközelítésére. Ezt az aspektust tudomásom szerint eddig nem vizsgálták. Részben érinti a témát Christopher Bernert 1983-as (a mindössze 30 éves szerző váratlan halála miatt posztumusz megjelent) írása – igaz, itt csak az amerikai szociológiáról írt, de megfigyelései nagyrészt általánosíthatók a szociológiára általában is. Ezt az áttekintést érdemes lenne kiterjeszteni az azóta eltelt 30 évre is. Erre természetesen itt nem vállalkozhatok, de néhány fordulópont megtalálására kísérletet teszek.

Bernert szerint világosan elkülöníthető szakaszok figyelhetők meg a szociológia története során a kauzalitás megközelítésében:

- a XIX. században végén a társadalomstatisztikai irányultságú szociológiában az okság kritika nélküli használata figyelhető meg, ezt váltja
- a XX. század elején a kortárs angol tudományfilozófia hatására, Karl Pearson-ban kikristályosodva a fogalom teljes elutasítása, ami
- a II. Világháború után a határterületek (biometrika, pszichometria, ökonometria, oktatástatisztika) felől érkező hatásokra egyfajta pragmatikus hasznosításra váltott, Lazarsfeld-del mint kiindulóponttal, Blalock-kal és Duncan-nel kicsúcsosodva – ők ismét okokról és hatásokról beszéltek.

Az amerikai szociológia Bernert által első generációnak nevezett iskolája (1890-1910), akárcsak az őt erősen befolyásoló német szociológia, társadalomstatisztikai irányultságú munkáiban kauzális terminológiát használt. Nagy mennyiségű, történelmi esettanulmányon alapuló adatot gyűjtöttek, ezekből kísérletek meg az okok meghatározását, aminek problémáját inkább a fogalmi szervezettségben, s nem a módszertanban látták. Empirikus módszertanuk John Stuart Mill induktív következtetésre vonatkozó szabályaira épült. Durkheimet kritikusan fogadták, részben éppen a szociológiai tények és az okságnak e tényekkel definiált megközelítése miatt. Ez a historikus tradíció teremtette meg a Columbia Egyetem empirikus reputációját.

A második generáció a Mill kritikájára épülő (W. S. Jevons-szal kezdődő és Karl Pearson-nal kikristályosodó) kortárs angol tudományfilozófia hatására került az empirikus eredmények kauzális interpretációját és általánosítását. Angliában Pearson körül erős statisztikai iskola intézményesült, szaklapokkal, egyetemi tanszékekkel. Az iskola Amerikában elsőként a Columbián, majd Chicagóban jelent meg. Elsősorban survey módszereket használtak, a kauzális terminológiát szinte egyáltalán nem alkalmazva. Beszédes, hogy az 1931-e, S. A. Rice szerkesztette *Methods in Social Science* a *causality* címszót csak a *historical studies* alatt tárgyalta. Az okság fogalmát a szerintük vele együtt járó törvényjellegű determinizmus miatt kritizálták. A paradigma alapmunkája Pearson 1911-es *The Grammar of Science*-e volt. Ez a nagyon olvasott mű expliciten a társadalomtudományoknak volt címezve, és évtizedekig meghatározta a statisztika oksággal kapcsolatos álláspontját. Pearson ebben egyértelműen elutasította az oksági kapcsolat fogalmának szükségességét, és elegendőnek tartotta csupán az együttjárással, a korrelációval helyettesíteni azt. Ez a szemlélet majd' száz évig tartotta magát, még az *Encyclopedia of Statistical Science* 1982-es kiadásában is csupán 2 oldalt szentelt a

causation-nek, míg tizenkét oldalt a *correlation*-nek (és talán még ma is hasonló arányokat találunk a szociológiai módszertan oktatási anyagaiban is). A hangnem érzékeltetése miatt érdemes ideidézni a *The Grammar of Science* XII. oldaláról a *Contingency and Correlation - The Insufficiency of Causation* cím alatt írt mondatot: „*Beyond such discarded fundamentals as 'matter' and 'force' lies still another amidst the inscrutable arcana of modern science, namely, the category of cause and effect.*” Pearson kifejti, hogy a kauzalitás empirikus jelenségből absztrahált dolog, nincs valódi létezése. Helyette a valószínűségi kapcsolatot hangsúlyozza, így jut el a kontingenciáig és a korrelációig, és ezek mérőszámaiig (lásd pl. Pearson-korreláció). Tehát a kauzalitás fogalmát félretéve, a neo-humeiánus megközelítéshez sorolhatóan együttjárások vizsgálatát, korrelációelemzést javasolt.

A harmadik generációt (1930-50) a határterületek (biometrika, statisztika, pszichometria, ökonometria, oktatástatisztika) felől érkező hatások határozták meg, az empirikus szociológiában Lazarsfeld tartozik ide. A kauzális terminológia még nem jelent meg újra, de a generáció előkészítette azt, ami végül a negyedik generációval (1950-70) történt meg. Ahogy Bernert találóan fogalmaz: az első generáció elméleteket alkotott módszer nélkül, a második módszertannal foglalkozott elmélet nélkül, a harmadik-negyedik generáció pedig azt vallotta, hogy e két dolog összeegyeztethető. Az elméletnek megfelelő módszereket, és empirikusan vizsgálható elméleteket kerestek. Statisztikai modelleket (faktorelemzés, útelemzés) használva okokról és okozatokról beszéltek. Blalock, Duncan fémjelzi a negyedik generációt (1960-1980), és Blalock 1964-es munkája, a *Causal inferences in non-experimental research*.

A Bernert által vizsgált utolsó időszakig, a '80-as évekig igaz volt az, hogy a kauzális következtetések a tudomány elméleti szintjéhez tartoztak, és az empirikus kutatás csak együttjárásokat és időbeli egymásutánokat tudott megalapozni. A robusztus összefüggés-alapú megközelítés volt egyeduralkodó, a kauzális állítások bíztak abban a feltételezésben, hogy minden releváns változó kontrollálva van. A kauzalitás egyöntetűen elfogadott meghatározása híján a fogalom leginkább a használati szabályok (asszociációs mérték, aszimmetria, kontrollálás) által voltak értelmezve. A kauzális terminológiát legexplicitebben használó módszertan a SEM (*structural equation modelling*) volt, melyet több kritika is ért David Freedman és mások részéről (lásd pl. Muthén, 1987) amiatt, hogy együttjárásokat (kovariancia-struktúrákat) értelmeznek oksági relációként.

A robusztus összefüggést ért kritikákkal kapcsolatban érdemes megemlíteni a két neves mobilitáskutató, Boudon és Hauser híres 1976-os vitáját is (Boudon, 1976, Hauser, 1976). Boudon a vitában Muthén fenti idézetére rímelve amellelt érvelt, hogy a Hauser által használt útmodellek paraméterbecslései nem szolgálhatnak oksági magyarázattal a változók kapcsolatát illetően, helyettük a háttérben működő mechanizmusokat magyarázó generatív modellekre lenne szükség. Boudon ezzel tulajdonképpen a mobilitáskutatás új, mikroszintű elméletének szükségességét hangoztatta. Boudon viszont olyan szimulációkat alkalmazott a státuselérés folyamatának modellezésére (státuszelérési folyamat egymást követő lépései különböző paraméterek függvényében, oktatási expanzió, szülők státusza stb. is paraméterezve a modellben), amelyek Hauser szerint leegyszerűsítőek, nem realisztikusak voltak. Boudon érvelése szerint ezeknek a modelleknek nem kell realisztikusnak lennie, mert célja nem a jelenségek leírása, hanem azok elméleti modellel való megmagyarázása. Modelljeit éppen

ezért generatív modelleknek nevezte, melyek a jelenséget generáló mechanizmusokat működés közben képesek megmutatni.

Hauser megközelítése tehát tulajdonképpen a neo-humeiánus irányzathoz sorolható, amely irányzat, a fenti említett viták tanúsága szerint, konfliktusos módon került szembe a 80-as, 90-es években körvonalazódó mechanizmus-alapú megközelítéssel. Ez a vita néhol (pl. Hedström és Swedberg, 1988) változó-alapú vs. mechanizmus-alapú megközelítések szembenállásaként fogalmazódott meg, a változó-szociológia (*variable-sociology*) elnevezésnek meglehetősen gúnyos felhangot adva.

Bernert úgy látja, hogy az általa azonosított történeti fordulópontokon általában más tudományokból jövő diffúziók okozták a szemléletváltozást. A mechanizmus-alapú megközelítés kialakulásában is erős volt a határtudományok (fizika, biológia, pszichológia, közgazdaságtan) felől érkező hatás. A módszertani individualizmus és a cselekvésemélet keretén belül, Merton és Coleman középszintű szociológiaelmélet-felfogásához csatlakozva a megközelítés egyes irányzatai (mint az analitikus szociológia) nem csak kutatómódszertani programként, hanem jól meghatározható szociológiaelméleti iskolaként is intézményesültek. A növekvő népszerűségű, ugyanezen megközelítéshez sorolható társadalmi hálózatkutatás jövőbeni perspektívái pedig a megközelítés további felfutását jósolhatják.

A mechanizmus-alapú megközelítés mellett az utóbbi két-három évtizedben a kísérletes módszerből kinövő tényellentétes és manipulatív megközelítés iskolája is megjelent, elsősorban az oksági megközelítés megfigyeléses vizsgálatokra történő (a megfelelő statisztikai módszerek kifejlesztése után lehetővé vált) kiterjesztésének köszönhetően. Az irányzat a szociológiában azonnal követőkre talált (pl. Berk, Newton, 1985, *propensity score matching*-gel a börtönbüntetés visszatartó erejéről), amihez a '90-es évek végén közgazdászok (pl. Heckman) is csatlakoztak közpolitikai programok értékelését lehetővé tevő technikák megalkotásával. A megfelelő adatgyűjtési (és elemzési) módszerekkel kapcsolatban élénk vita bontakozott az irányzaton belül az ezredforduló körül; a vita lényegében a kontrollált kísérlet és a megfigyeléses adatokon alapuló strukturális egyenletek modellezés (SEM) szembeállításáról bontakozott ki. Jó összefoglaló az érvekkel kapcsolatban Heckman (2005) cikke, Michael E. Sobelnek erre adott kommentára, majd Heckman viszontválasza a *Sociological Methodology*-ban. Míg Heckman a kontrollált kísérletek közpolitikai alkalmazásának egyeduralma ellen érvelt (szerinte azok az igazán releváns kérdésekre nem tudnak választ adni), a megfigyeléses adatokon alapuló SEM-modellezést ajánlva helyettük, addig Sobel a SEM kauzális interpretálhatóságát kérdőjelezte meg. Az utóbbi években a tényellentétes megközelítés további terjedése és népszerűségének növekedése figyelhető meg a szociológiában (lásd pl. Morgan és Winship 2007-es nagy összefoglaló munkáját).

1.7. Mai tendenciák: Big Data – szükségtelen kauzalitás?

A szociológiatörténeti áttekintést itt befejeztem volna, ha Csepeli György nem hívja fel a figyelmet a kauzalitás fogalmának ismételt elutasítására, ezúttal a Big Data²² oldaláról. Ezért egy rövid bekezdésben, csak néhány gondolat erejéig kitérnék a korreláció-kauzalitás dichotómia ezen újabb felfutására, a szociológiát ismét a határterületek (informatika, statisztikus fizika, matematika) felől érő kihívásra.

A probléma gyökere a Big Data adatgyűjtésének a megszokottól eltérő módjában van: nem survey-alapú, hanem teljeskörű adatok állnak rendelkezésre, melyekhez nem az egyének megkérdezésével, hanem azok mindennapi viselkedésének digitális rögzítésével jutunk. Mindkét jellegzetességet előnyként szokták bemutatni az irányzat képviselői. A teljeskörűség kétségtelenül nagy előny, ugyanakkor, mivel ezek az adatok mintegy melléktermékként jönnek létre, az adatokból az információhoz vezető út további inputot (elméletet, háttértudást) igényel. A legjobb algoritmus sem képes a jelentéktelen adatok/korrelációk megkülönböztetésére a fontosaktól.

A másik jellegzetesség, a viselkedés digitális rögzítése viszont nem problémátlan. Jó példa erre a boldogságmérő, mellyel (lásd hedonometer.org tudományos publikációit) szövegek, pl. Twitter-szövegek boldogság-szintjét mérik alkotószavaik boldogságtartalmának átlaga alapján. Az egyes szavak boldogságtartalma a kontextustól függetlenül adott, tesztalanyok korábbi pontozása alapján, pl. a *hope*, *hero*, *to win* magas pontszámmal rendelkeznek. A kutatók szerint ezzel populációk, földrajzi régiók aktuális boldogságát objektívebben tudjuk mérni, mint a hagyományos, szubjektív elégedettség-re rákérdező survey-ekkel. Az objektivitással kapcsolatos problémák, azt hiszem, nyilvánvalóak, itt nem részletezném őket - elég csak a megformált szöveg kulturális beágyazottságára utalni. Más problémák mellett éppen a kulturális kontextus (a stílus, a nyelvi normák) figyelmen kívül hagyása, a szövegnek egyfajta információközlő viselkedéssel való azonosítása miatt nehéz a kutatók eredményeinek (pl. az életkor-boldogság, vagy az Egyenlítő-től való távolság-boldogság összefüggésének) interpretálása.

A Big Data-információkat elsősorban (gazdasági célú) előrejelzésekre használják, s ritkán merül fel a megfigyelt mintázatok magyarázatának kérdése. Ahogyan egy 2013-as Big Data-bestseller (Mayer-Schönberger, Cukier) fogalmaz: a miért-re vonatkozó kérdéseket a mi-re vonatkozókra kell cserélni. A korrelációelemzés a kísérleteken alapuló oksági elemzéssel szemben gyors és olcsó, ráadásul „világos betekintést ad, ami újra zavarossá válhat, amint az okságot visszahozzuk a képbe”. Hogy e gondolatok kontextusát lássuk: példája szerint az, hogy a narancssárga autók adatbányászati eredmények szerint ritkábban mennek tönkre, a használtautó-kereskedők számára, előrejelzésként értékes információt jelent, mégsem érdemes sem a miért-en gondolkodni, sem narancssárgára festeni autónkat. Chris Andersonnak, a

²² A Big Data-t szándékosan nem fordítom magyarra. A fogalom, ahogy látom, egyszerre fedí a világszerte napi szinten előállított óriási digitális adatmennyiséget, ennek kezelési-elemzési problematikáját, információs potenciáját, és az utóbbi ma még beláthatatlan gazdasági-társadalmi hatásait is.

Wired főszerkesztőjének 2008-as provokatív cikke²³ ugyanerről: „A hatalmas adattömegek elérhetősége és az elemzésükre alkalmas statisztikai eszközök a világ megértésének új útját kínálják. A korreláció fontosabbá válik az okságnál, és a tudomány anélkül is haladást tud elérni, hogy koherens modellekre, egységes elméletekre vagy mechanisztikus magyarázatokra támaszkodna.”

Úgy tűnik, a tudománytörténet fejezetei ismétlik önmagukat. Az okságot ismét a korreláció felől éri kihívás, mint Pearson idején. A kihívásra felhozott, kauzalitás melletti érvek is ismerősek lehetnek, akár egy szociológia alapszakos diáknak is. Például²⁴: a korreláció megtalálása nem a kutatás végpontjaként, hanem inkább hipotézisek kiindulópontjaként lehet fontos. A Big Data olyan rendszerek esetén használható előrejelzésre, amelyek időben konzisztensek, nem változnak megjósolhatatlanul, kis komplexitásúak – a narancssárga autókat termelő autóipar ilyen lehet, a társadalmi rendszerek általában nem ilyenek. A tudományos elméleteknek nem az előrejelzés a végső célja, hanem a magyarázat, ami viszont oksági elméletet kíván. A legjobb algoritmus sem képes értelmes kérdésfeltevésre, a mérési adatok interpretálására vagy helyes konklúziók levonására. A mérési adatok (tweetek) nem azonosak a mért valósággal (boldogság). A fontos korrelációk megkülönböztetése a jelentéktelenektől oksági modellt kíván. Stb.

A társadalmi Big Data kutatói ismét a természet- és bölcsészettudományok közötti határvonalon, a tudományos objektivitást érintő episztemológiai kérdések kereszttüzében találják magát. A Big Data kétségtelenül korszakváltást jelent a szociológia számára. A hatékony algoritmusokon kívül azonban, ahogy a felmerülő kérdések kapcsán látható, szükség van a szociológia és a szociológiai módszertan (újra)felfedezésére is. Így explicitté válhatnának a szociológia számára ismerős olyan interpretációs csapdák, mint a hamis korreláció lehetősége (tweetek boldogsága - házasságok aránya), a hatásirány kérdése (Twitter-hálózatban csúcsok fokszáma és tweetjeinek boldogsága összefügg, de nem csak az előbbi lehet az ok és utóbbi az okozat) vagy az elemzési szintek meg nem különböztetéséből fakadó ökológiai tévkövetkeztetés lehetősége (az átlagosan iskolázottabb államokban magasabb a boldogság átlagos szintje). Sőt továbblépve, ahogy Szántó és SYI írják kitűnő, 2010-es cikkükben, a megfigyelt mintázatok megértése, az érdekek, szándékok és normák figyelembevétele is felmerülhetne. Az USA államaiban a boldogságmérő karácsony körül a boldogság rendszerszerű emelkedését mutatja. Az ünnep által kiváltott tényleges öröm vagy a jókívánságokra vonatkozó, tweetekre is érvényes társadalmi konvenciók állnak-e emögött?

4. Megfigyeléses vizsgálatok a Rubin-Holland modell keretében: *confounding*

A Rubin-Holland modellt sok kritika (lásd pl. a Holland, 1986 cikk hozzászólásait) érte amiatt, hogy „felesleges absztraktumokat” vezet be és tárgyal. Az alábbiakban cáfolni

²³ http://www.wired.com/science/discoveries/magazine/16-07/pb_theory

²⁴ Az érvek egy része Gil Press-nek a forbes.com-on közzölt publicisztikájából származik, ahol a Big Data-szkeptikusak érveit gyűjti össze (<http://www.forbes.com/sites/gilpress/2013/03/11/whats-to-be-done-about-big-data/>).

szeretném ezt az álláspontot. Megpróbálom megmutatni, hogy a modell a kauzalitás fogalmának világos definíciójával segít a szociológiában általános megfigyeléses vizsgálatokkal kapcsolatos szemlélet alakításában is. A *confounding* fogalmát tisztázom majd, a Rubin-Holland modell keretén belül, Heckman és Robb (1985) közgazdaságtani hatásvizsgálatokkal kapcsolatos eredményeit és Judea Pearl szemléletét is hozzákapcsolva a modellhez. A két előbbi megközelítést Xie (2013) is összekapcsolja egy friss tanulmányban; a jelen fejezetben e publikációk folytonos függő változóra (számértékű hatásmutatóra) kidolgozott eredményeit általánosítom bináris függő változóra (gyógyult-e a beteg, sikeres volt-e a felvételi stb). Megmutatom a szemlélet előnyét más tényellentétes-alapú munkával (Greenland és Robins, 1986, 2009) szemben. Pearl és a Rubin-Holland modell tudásom szerint máshol nem szerepel közös keretben, itt jól látjuk majd, hogyan egészítik ki egymást.

Összefoglalva a Rubin-Holland modellnek a megfigyeléses vizsgálatokra vonatkozó következményeit: ha a hatást az u egyedre az $Y_t(u)-Y_c(u)$ különbséggel definiálom, akkor az Oksági Következtetés Fundamentális Problémájába ütközöm: lehetetlen ugyanazon egyeden a kezelés és a kontroll csoportba sorolás esetén mért értéket is megismerni. A probléma a definíció tényellentétes voltából fakad.

A fundamentális problémának az empirikus szociológiában is alkalmazott statisztikai megoldása: nem az egyénenkénti, hanem csak az átlagos hatást, a

$$T := E(Y_t) - E(Y_c)$$

különbséget mérjük. Bár a fundamentális probléma miatt ez sem megfigyelhető, a

$$T_{PF} = E(Y_t | S=t) - E(Y_c | S=c)$$

prima facie átlagos hatás meg tudjuk figyelni (lásd 2. ábra).

Ha nem megfigyeléses vizsgálatunk van, hanem randomizált kísérletet végzünk (a kezelt/kontroll allokációt random módon végezzük), akkor a randomítás folytán a kezelt/kontroll kiosztás nyilván az Y válaszoktól is független:

$$(Y_t, Y_c) \perp S,$$

Ebből

$$E(Y_t | S=t) = E(Y_t)$$

és

$$E(Y_c | S=c) = E(Y_c),$$

amiből következően T és T_{PF} egyenlő. Ám megfigyeléses vizsgálat esetén az allokáció nem random. A szociológia számára legtöbbször megfigyeléses adatok állnak rendelkezésre, mert a random allokáció kivitelezhetetlen vagy etikátlan lenne. Az allokáció ilyenkor legtöbbször *informatív*²⁵, vagyis önmagában az is a vizsgálat kimenetét érintő információt hordoz, hogy a kezelt vagy a kontroll csoportba sorolódott (vagy sorolta magát) a megfigyelt alany.

²⁵ Az *informatív allokáció* kifejezés Rudas (2010) találó megfogalmazása.

Érdemes ebből a szempontból végiggondolni a dolgozat elején megadott néhány típuskérdést, azt feltételezve, hogy megfigyeléses adatokból szeretnénk rájuk választ adni.

Mekkora munkaerőpiaci előnyt biztosít a diploma? Tegyük fel, hogy a megfigyeléses adatok az ugyanazon évben érettségizett diplomások és nem diplomások keresetét tartalmazzák. Itt egyrészt az egyének sorolják be magukat a kezelt/kontroll csoportba (döntve arról, hogy jelentkeznek-e egyetemre), másrészt a felvételi eljárás szelektál bizonyos érdemek szerint.

Eredményes volt-e az OFA munkáltatóknak kiírt 2009-es foglalkoztatottság-megőrző programja? Ismét kétlépéses a szelekció: az önszelekció során a cégek döntenek a pályázásról, majd az elbírálás során a pályázók közül néhányan „érdemeik” alapján kerülnek a nyertesek közé.

Hat-e a jövedelmi helyzet a politikai részvételre? Az egyes jövedelmi csoportokba kerülés nyilván a legkevésbé sem random, az alokációt meghatározza az iskolai végzettség, a foglalkozási szektor, az egyéni rátermettség stb. A felsorolt tényezők mindegyike összefüggésben állhat a politikai részvétellel, tehát $Y_t, Y_c \not\perp S$. Ezek a tényezők az adott hatásmechanizmusban *confounder* szerepet játszhatnak, az 1.5.1 fejezetben definiált értelemben: bevonásuk után megváltoztathatják (eltüntethetik vagy ellenkező irányúra fordíthatják) a mért nyers hatást.

Diszkriminálnak-e a munkaadók a roma álláskeresőkkel szemben? Tegyük fel, hogy megfigyeléses adatok állnak rendelkezésre adott állásokra jelentkező roma és nem roma jelentkezők sikerességéről. A romaság nyilván nem random módon oszlik meg a társadalmi rétegeken belül, az iskolázatlanok között több a roma, és az iskolázottság összefüggésben áll az álláskereső sikerével is, vagyis ismét: $Y_t, Y_c \not\perp S$. Az iskola az adott helyzetben potenciális *confounder*.

A randomitás hiánya esetén az erős ignorálhatóság feltevésével szabadulhatunk meg a fundamentális problémától. Ha tudunk néhány olyan háttérváltozót (X) azonosítani, amikre feltéve a kezelt/kontroll kiosztás (S) már független lesz a potenciális választól (Y_t -től és Y_c -től):

$$(Y_t, Y_c) \perp S \mid X, \quad (3)$$

akkor az X -re korrigált T és T_{PF} egyenlő. A korrigálásnak különböző statisztikai megoldásai vannak (az X szerint rétegzést végezhetünk, vagy a kezelt/kontroll alanyokat X vagy – mint a *propensity score matching*-nél – X egy megfelelő függvénye szerint párosíthatjuk, vagy az X -et, esetleg annak valamely megfelelő függvényét független változóként szerepeltethetjük regressziós modellben). A fenti példák esetén felsorolt potenciális *confounderek* az X elemeiként jöhetnek szóba, pl. kérdés, elég-e az álláskereső iskolázottságára kontrollálni ahhoz, hogy a romákkal szembeni diszkriminációt bizonyíthassuk.

A Rubin-Holland modell definícióból adódó másik feltevése (*Stable Unit Treatment Value Assumption, SUTVA*), hogy a kezelés adott egyeden kimért hatása független attól, hogy más egyedek megkapták-e a kezelést. Ez társadalomtudományi vizsgálatokban nem mindig

teljesülő feltevés, s mint ahogy a 1.5.4. fejezetben említettük, az adatgyűjtés során kell erőfeszítéseket tenni legalább részleges elérésére.

Tehát: a SUTVA és a fundamentális probléma következtében elkerülhetetlen további (általában nehezen tesztelhető) feltételeknek a bevonása a modellbe. A Rubin-Holland modell lényege a megfigyeléses vizsgálatok esetén éppen az, hogy ezek a feltételek explicitté válnak, és hangsúly kerül legalább részleges tesztelésük szükségességére. Így a modell egyik előnye éppen a T és T_{PF} eltérésének és a *confounding*-nak a világos értelmezése. Vegyük ugyanis észre, hogy a T_{PF} tulajdonképpen azonosítható azzal az asszociációval is, amit a robusztus összefüggés-megközelítés az analízis további lépéseiben újabb változók (*confounderek*) bevonásával kontrollálni igyekszik. Vagyis a tény, hogy a megfigyelt asszociáció nem feltétlenül azonosítható a tényleges hatással, a robusztus összefüggés-megközelítésben is nyilvánvaló, még ha ott az utóbbi nincs is pontosan definiálva.

Az előbb említett *confounding* a megfigyeléses vizsgálatok alapfogalma, a következőkben ennek megragadásáról szeretnék írni a Rubin-Holland modell nyújtotta kereten belül. A *confounding* jelenségét John Stuart Mill tárgyalta elsőként induktív logikai vizsgálataiban, majd Yule 1930-ban, már explicit statisztikai fogalmakkal, kísérletek elemzésével kapcsolatban. Ő *fictitious association*-ként említi azt a jelenséget, amikor a kísérlet eredménye nemként külön elemezve mást ad, mint aggregálva. Részletesebben lásd Greenland és Robins (2009) történeti összefoglalóját. Mára a jelenség az orvosi és szociológiai alkalmazásokban széles körben ismertté vált, a szociológiaoktatásban használt jegyzetek alappéldája (pl. Freedman, Pisani, Purves: Statisztika, Typotex). A *confounding* néhol hamis (*spurious*), máshol képzelt asszociációként (*fictitious association*) szerepel. Mindegyik kifejezés arra utal, hogy a megfigyelt nyers asszociációt érvénytelennek tartjuk, és azt gondoljuk, hogy a valódi asszociáció eltér tőle. Ez a megfogalmazás így nagyon alulhatározott. Egyrészt a „valódi asszociáció” nincs definiálva - mint látni fogjuk, itt éppen az adja a Rubin-Holland modell hasznát, hogy pontos definíciót ad rá. Másrészt nem világos, hogy miért gondoljuk a feltételes (pl. nemként számolt) hatást érvényesebbnek, mint a nyers hatást – vagyis nem világos, hogy a *confounding* valóban mindig hamis asszociációt mutat-e. Erre a következő fejezetben térek ki.

A kifejezést szándékosan nem fordítom magyarra – nincs rá bevett fordítás²⁶, általában az angol alakot szerepelteti a magyar szakirodalom. Több, nem teljesen átfedő jelentésben használt (Greenland, Robins, 1986). Egyrészt a kezelt és kontroll csoport inherens eltéréséből adódó közvetlen össze nem vethetőségét értik alatta (*noncomparability*). Másrészt az asszociációs mérőszámok össze nem vonhatóságát (*noncollapsibility*): vagyis valamely harmadik változóra vett feltételes rétegekben kapott asszociációs mérőszámok eltérését a nyers, nem rétegzett mérőszámtól. Van egy harmadik, szintén bevett definíció is (pl. Miettinen, Cook, 1981.), mely szerint a *confounder* az okkal (itt: allokációval) és az okozattal is kapcsolatot mutat. Ez tulajdonképpen ekvivalens a *noncollapsibility*-vel. Az utóbbi

²⁶ Kende Gábor és Szaitz Mariann özszemosódásként fordítja (Freedman, Pisani, Purves: Statisztika, Typotex), de ez a magyarítás még nem elterjedt.

definíciót Judea Pearl (pl. 2014) azzal egészíti ki, hogy a hatásirány olyan, hogy a *confounder* hat az okra és az okozatra is (erről részletesebben a következő fejezetben írok). Mint a következő fejezetben látni fogjuk, ez a feltétel azért szükséges, mert a *noncollapsibility* fennállása általában nem vonja maga után automatikusan a nyers asszociáció érvénytelenségét, vagyis a *confounder*re történő kontrollálás szükségességét. Vegyük észre, hogy az első definíció lényegesen különbözik az utóbbiaktól: a *noncomparability*-hez nem feltétlenül kell tudni beazonosítani egy harmadik, *confounder* változót.

Az 1. táblázat a *noncollapsibility* egy példáját mutatja; a számok valós adatok egy vesekő műtéti terápiás vizsgálatából (Julious, Mullee, 1994). Az első táblázat szerint a nyílt műtét a kevésbé hatékony műtéti technika, mert az így műtött betegek 5%-kal kisebb arányban (háromnegyed akkora eséllyel) gyógyultak. Mindkét mérőszám bevett asszociációs mérték, az arányok különbsége az epidemiológiában kockázat-különbségként (*risk difference*) ismert, az esélyhányadost is széles körben használják.

Ugyanakkor a harmadik változó, a vesekő mérete szerinti rétegekben a nyílt műtét tűnik preferálандónak: 4, ill. 6%-kal nagyobb arányban; 23, ill. 108%-kal nagyobb eséllyel gyógyultak az így műtött betegek. A *noncollapsibility* fennáll: a rétegzett elemzésből adódó asszociációs mérőszámok lényegesen különböznek az aggregált tábla mérőszámától.

Sőt, a példa valójában a *noncollapsibility* Simpson-paradoxonként ismert esete: az asszociációs mérőszám a rétegzett elemzésben ellenkező előjelet (!) mutat, mint az aggregált táblán. A paradoxon magyarázata az allokáció informativitása: a nagyobb köveket inkább nyílt műtéthez rendelik az orvosok. A kő nagysága azért informatív jelen esetben a vizsgálat kimenetére nézve is, mert a nagyobb kövek gyógyulási mutatói rosszabbak a kisebb kövekénél (bármelyik műtéti eljárást alkalmazzák is). Vagyis az aggregált adatok azért szólnak a PCNL mellett, mert a PCNL-lel kezelt csoportban a könnyebb esetnek minősülő kis-köves betegek voltak.

Hasonló jelenség több valós vizsgálatból ismert, pl. a Berkley felvételi statisztikáival kapcsolatban (Bickel et al., 1975). A táblázat fejléceit megkettőzve ezt a kerettörténetet is leírhatjuk (nem a Berkley eredeti adataival, de annak lényegét követve). A műtét sikerességét a felvételi sikerességére, a műtét típusát a felvételiző nemére, a kő nagyságát pedig a megcélzott szakra való bejutás nehézségére cserélve új kerettörténet bontakozik ki. Az aggregált tábla szerint a nők felvételi esélyei rosszabbak a férfiakénál, de a szak nehézségére kontrollálva ez a következtetés az ellenkezőjére fordul: a nők nagyobb eséllyel jutnak be mindkét szaktípusra. Magyarázat: bár a nők felvételi mutatóik kedvezőbbek a férfiakénál (bármely szakról van is szó), az aggregált adatokból ez nem látszik, mert a nők nehezebb szakokra felvételiznek (vagy a nők által preferált szakokon nehezebb a felvételi, ahogy Rudas 2010 fogalmaz, a szokásos interpretációval szemben meghagyva a nőkkel szembeni diszkrimináció lehetőségét is). A nemethr.web.elte.hu/causality.xlsx file non-collapsibility c. munkalapján megtalálhatók ezek a táblázatok, ahol az adatok megváltoztatásával megfigyelhető, hogyan változnak a hatásmutatók, milyen együttállások esetén alakul ki és mikor nem a Simpson-paradoxon.

Összesen	Felvett (Sikeres műtét)	Fel nem vett (Sikertelen műtét)	Összesen
Nő (Nyílt műtét)	273	77	350
	78%	22%	100%
Férfi (PCNL)	289	61	350
	83%	17%	100%
A nők (a nyílt műtét) hátránya a férfiakhoz (PCNL-hez) képest: sikerességi arány különbsége -5%, esélyhányados 0,75			

Nehéz szak (Nagy vesekő)	Felvett (Sikeres műtét)	Fel nem vett (Sikertelen műtét)	Összesen
Nő (Nyílt műtét)	192	71	263
	73%	27%	100%
Férfi (PCNL)	55	25	80
	69%	31%	100%
A nők (a nyílt műtét) hátránya a férfiakhoz (PCNL-hez) képest: sikerességi arány különbsége 4%, esélyhányados 1,23			

Könnyű szak (Kis vesekő)	Felvett (Sikeres műtét)	Fel nem vett (Sikertelen műtét)	Összesen
Nő (Nyílt műtét)	81	6	87
	93%	7%	100%
Férfi (PCNL)	234	36	270
	87%	13%	100%
A nők (a nyílt műtét) hátránya a férfiakhoz (PCNL-hez) képest: sikerességi arány különbsége 6%, esélyhányados 2,08			

1. táblázat A *confounding* egy példája: az asszociációs mérőszámok össze nem vonhatósága. Vastagítva az egyetemi felvételi, normál betűtípussal a vesekő-műtési kerettörténettel.

Ahogy Greenland és Robins (1986) megjegyzi, itt a *confounding* előállása függ attól, hogyan definiáljuk a kapcsolati mutatót. Pl. mutathat *confounding*ot az esélyhányados, míg az aránykülönbség nem. Rudas (2010) tárgyalja általánosan a hatásmutató megválasztásának és a *collapsibility* összefüggésének problémáját. A *collapsibility* és a *confounding* összefüggését máshol (pl. az 1.5.1 fejezetben definiált értelemben) nem kísérleti/kezelt összevetéssel vezetik be, hanem általánosabban: a fogalom ilyenkor arra utal, hogy egy harmadik változó szerint kontroll megváltoztatja bármely két változó közötti kapcsolat előjelét/nagyságát. Egymás után vagy akár egy lépésben több változóra is kontrollálhatunk, ezek mindegyike szerepelhet *confounder*-ként, ilyenkor tulajdonképpen a *confounder* szerep függ attól, hogy aktuális mely változókra kontrolláltunk már. Fontos rámutatni, hogy a *confounding* utóbbi értelmében a „valódi hatás” nehezen definiálható, nem világos; impliciten szokás azt is beleértetni a fogalomba (lásd a robusztus összefüggés megközelítést), hogy a (minél több mért változóval) kontrollált kapcsolat mutatja a valódi hatást, vagy legalábbis az van lehető legközelebb hozzá. Ezzel szemben a *confounding*-nak a következőkben általam használt, tényellentétre épülő értelmében a valódi hatás, *T* egyértelműen definiált, és a *confounding* is ehhez a definícióhoz képest értelmezendő.

A Rubin-Holland modell keretére térve: Greenland és Robins (Greenland, Robins, 1986, és Greenland, Robins, 2009) kapcsolta elsőként össze a *confounding*ot a tényellentétes megközelítéssel (bár nem expliciten a tényellentétes fogalmával kapcsolatban, és Rubin és Holland munkáját sem idézik, mert, mint a 2009-es cikkben írják, nem ismerték). Ők a *confounding*ot a kontroll- és a kezelt csoport fel nem cserélhetőségével azonosítják (*noncomparability*). Formális definíciójuk megértéséhez szükséges lesz néhány jelölés bevezetésére. Aszerint, hogy a vizsgálati alanyok hogyan reagálnának a kezelésre és a kontrollcsoportban való részvételre, az alanyok négy típusát különböztetik meg (2. táblázat). A tényellentétes megközelítés jól látható a feltételes módból („reagálnának”): nem tudhatjuk, hogyan reagálnának mindkét esetben, vagy csak kezelt- vagy csak kontrollcsoportba sorolva ismerhetjük meg reakciójukat. A négy csoport: (1) az immunisak (akik egyik csoportba sorolva sem betegednének meg), (2) azok, akiket a kezelés meggyógyítana, (3) azok, akiket a kezelés beteggé tenné, végül (4) azok, akik reménytelenek (mindkét csoportban megbetegednének). Tegyük fel, hogy megtörtént az allokáció. Jelölje p_i , $i=1..4$ a kezeltbe allokáltakon belül az egyes típusok arányát, ugyanezen arányok legyenek q_i -vel jelölve a kontrollba allokáltakra. Ekkor a két csoport összevethetőségét, vagyis a *confounding* hiányát Greenland és Robins meghatározásában a két csoportban a betegségtől menteseknek a kezelés hiánya esetén adódó arányának azonossága, vagyis a

$$p_1 + p_3 = q_1 + q_3 \quad (4)$$

összefüggés definiálja.

Alany típusa	A típus definíciója	Kezeltként	Kontrollként
1. típus	Immunis eset	1 (egészséges)	1
2. típus	A kezelés gyógyít	1	0 (beteg)
3. típus	A kezelés káros	0	1
4. típus	Reménytelen eset	0	0

2. táblázat A Greenland-Robins tényellentétes megközelítés

A szerzők maguk is megemlítik, hogy az összevethetőségnek ez egy megengedő definíciója (részleges összevethetőség, *partial comparability*), és hogy szigorú értelemben a teljes összevethetőség (*full comparability*) megkövetelné azt is, hogy kezelés hatására is azonos arányok adódjanak a betegségtől mentesekre a két csoportban, vagyis a

$$p_1 + p_2 = q_1 + q_2 \quad (5)$$

összefüggés teljesülését. Azonban nem indokolják, miért tartják a megengedőbb feltétellel is elkerülhetőnek a *confounding*-ot. Azt gondolom, talán az orvosi alkalmazásokban gyakori követéses vizsgálatokból indultak ki, ahol a két csoport kiinduló állapotbeli azonossága, a *baseline* különbség hiánya van a kísérlettervezés fókuszában, és az valóban így definiálható. Magam ezt a megengedő definíciót problémásnak látom, a későbbiekben a problémát egy általánosabb keretben részletezem.

A részleges összevethetőség (4)-beli definícióját nevezi tehát az epidemiológiai irodalom a kezelt és a kontroll közötti *baseline* különbség hiányának. A vizsgálatok tervezésénél ennek elérésére bizonyos háttérinformációk birtokában ténylegesen törekedni is szoktak, a kezelt és a kontroll közötti, a *baseline* túléléssel kapcsolatban levő, ismert faktorok szerint kiegyensúlyozott allokációval.

Greenland és Robins (más fogalmakkal, más jelöléssel, a tényleges hatást expliciten nem definiálva) a T_{PF} és T eltérésének problémájára is kitér. Nem a Rubin-Holland modellben dolgozva, de a probléma azzal ekvivalens formalizálásával *non-identifiability*-ként hivatkozzák azt, hogy a megfigyelt hatásból a tényleges hatás nem megítélhető. A megfigyelt hatást a betegségtől mentesek arányának a kezelt- és a kontrollcsoport közötti eltéréssel definiálhatjuk, a fenti jelölésekkel:

$$T_{PF} = (p_1 + p_2) - (q_1 + q_3)$$

A T_{PF} -ből nem tudunk következtetni a valódi hatásra (pl. p_2 és p_3 értékére), még az előjeléből vagy 0 voltából sem. Ha most feltesszük, hogy a vizsgálattervezés során sikerült a részleges összevethetőséget, vagyis a $p_1 + p_3 = q_1 + q_3$ feltétel teljesülését biztosítani, akkor az előbbi egyenlet a

$$T_{PF} = p_2 - p_3$$

alakra egyszerűsödik. Ahogy Greenland és Robins felhívja rá a figyelmet: ekkor már informatív a T_{PF} előjele: pozitivitása esetén pl. p_2 pozitivitására, vagyis arra következtethetünk, hogy néhány alanynál a kezelés gyógyhatású. Hasonlóan, negatív előjel a p_3 pozitivitására, vagyis néhány alanynál a kezelés hatására bekövetkező romlásra következtethetünk. Vagyis egy feltétel bevonásával közelebb kerültünk a valódi hatás megismeréséhez, de nehézkesen értelmezhető módon. A továbbiakban ismertetett keretben a tényleges és a megfigyelt hatás közötti különbség egy kézenfekvőbb megközelítését mutatom be.

A fenti megközelítés a p_i és q_i paraméterekkel paraméterezte az eloszlást. A megközelítés nehézségét éppen az okozza, hogy ezek közül egyik paraméter sem adható meg a megfigyelt adatokból. Az alábbiakban a T és a T_{PF} közötti eltérést hiányzó adat problémaként definiáljuk, úgy, hogy minimális számú paraméter legyen nem megadható a megfigyelt adatokból. Az egyszerűség kedvéért a hatást adott minta mellett határozzuk meg, nem foglalkozva mintavételi, külső validitási, ill. a mintavétel ismétlődése esetén adódó, várható értékkel megközelíthető problémákkal.

A tényellentétes megközelítés szükségessé teszi egy olyan operátor bevezetését, ami egy manipuláció eredményét jelöli. Az 1.5.4. fejezetben, bár világosabbá tette volna a jelölést, nem említettem ezt a lehetőséget, mert ott a Rubin-Holland modell eredeti formalizmusát követtem. A klasszikus statisztikai formalizmusban nem szerepel ilyen operátor, mert ott szokásos módon a megfigyelt együttes eloszlásból indulunk ki, nem foglalkozunk nem-megfigyelt eseményekkel. A tényellentétes okság mesterséges intelligenciakutatásban alkalmazott iskolájának legismertebb képviselője, Judea Pearl (pl. Pearl, 2009) jelölését követve a bevezetett operátort do-operátorként fogom jelölni. Pl. a

$$E(Y \mid do(S=t))$$

formula az Y várható értékét jelöli azon feltétel mellett, hogy mindenkin elvégeztük a kezelést. Az 1.5.4. fejezetben ezt a kevésbé kézenfekvő $E(Y_t)$ -vel jelöltük. Vegyük észre, hogy ez az érték különbözik az

$$E(Y \mid S=t)$$

értéktől: utóbbi az $S=t$ esetén megfigyelt Y várható értékét fedi, ami egy aktuális együttes eloszlás, egy aktuális allokáció függvénye. Továbbá, ha Z egy valószínűségi változó, akkor az

$$E(Y \mid do(S=t), Z=z_0)$$

kifejezés jelentése: Y várható értéke, ha mindenkin, akinek Z szerinti megfigyelt értéke z_0 , elvégezzük a kezelést. A pontosság érdekében különbséget fogunk tenni az allokáció (A) eredménye ($A=1$ treatment-csoport, $A=0$ kezelt-csoport, A értékét rögzítettnek tekintjük, nem fogjuk manipulálni) és a kezelés tényleges megtörténte ($S=t$ vagy c , mint a Rubin-modellben, S értékét potenciálisan változtathatónak tekintjük) között.

Most már világosan definiálhatjuk a T és a T_{PF} közötti eltérést. A T átlagos hatás bináris válasz ($Y=1$, ha betegségtől mentes, 0 , ha beteg az alany, mintegy hasznát, *utility*-t rendelve az egyes állapotokhoz²⁷) esetén az egészségesek arányának azon két lehetséges világban mért értéke közti különbséggel definiálható, amikor a két világot a mindenki kezelése ill. a senkinek nem kezelése adja:

$$T = E(Y \mid do(S=t)) - E(Y \mid do(S=c)) = \frac{n_t}{n} - \frac{n_c}{n},$$

ahol n_t és n_c az alábbi táblázatból jön (n a mintanagyság, n_t gyógyultak száma, ha mindenkit kezelünk, n_c gyógyultak száma, ha senkit sem kezelünk):

	<i>Egészséges</i>	<i>Beteg</i>
Kezelés mindenkin	n_t	$n - n_t$
Kontroll mindenkin	n_c	$n - n_c$

3. táblázat A tényleges hatást leíró tábla, a cellákban esetszámok

A tábla félkövérrel szedett értékeit nem ismerjük, a sormarginálisok ismertek, homogének.

²⁷ A bináris változó (gyógyult/beteg) kódolását lényegében önkényesen választottuk fent 1/0-nak. Ha más kódolást (tetszőleges $a > b$) választunk (értsd: a gyógyulásra vonatkozóan más hasznossági értéket), akkor a T_{PF} hatás előjele változatlan, de értéke változik (hisz a kódolás változása éppen a gyógyulás hasznát skálázza át). A T_{PF} változása, ha d volt az 1/0 kódolás mellett kapott T_{PF} értéke, $d(a-b)$ lesz, tehát az eredeti hatás szorozódik a gyógyulás hasznával, mintegy járulékos haszonnövekedést jelezve.

A T_{PF} hatást definiáló megfigyelt tábla (n_1 és n_0 a treatmentbe/kontrollba alokáltak száma, n_{11} a treatmentbe alokáltak közül a gyógyultak száma, ha mindenkit kezelünk közülük, n_{c0} a kontrollba alokáltak közül a gyógyultak száma, ha senkit nem kezelünk közülük):

	<i>Egészséges</i>	<i>Beteg</i>
Kezelés kezeltbe soroltakra	n_{11}	$n_1 - n_{11}$
Kontroll kontrollba soroltakon	n_{c0}	$n_0 - n_{c0}$

4. táblázat A megfigyelt tábla

A megfigyelt hatás az egészségesek arányának két csoport közti különbségével definiálható:

$$T_{PF} = E(Y | do(S=t), A=1) - E(Y | do(S=c), A=0) = \frac{n_{11}}{n_1} - \frac{n_{c0}}{n_0}.$$

Ezt a mutatót nevezik az epidemiológiában a kísérleti elrendezéstől függően prevalencia-különbségnek (keresztmetszeti vizsgálat), vagy járulékos kockázatnak (*relative risk*) ill. kumulatív incidencia-különbségnek (követéses, un. zárt kohorsz-vizsgálat). Bár tudjuk, hogy más hatásmutató is lehetséges volna (pl. esélyhányados), a hatásmutató megválasztása itt nem önkényes, hisz közvetlenül adódik abból, hogy bináris *outcome*-ra alkalmazzuk Rubin és Holland definícióját. Ugyanakkor a választásnak a mutató tulajdonságaiból adódó következményei vannak. Gondoljunk arra a korábbi megjegyzésre, hogy a *confounding* (vagy a Simpson-paradoxon) előállása függ a kapcsolati mutató megválasztásától. Rudas (2010) általánosan vizsgálja azt a kérdést, hogy milyen mutató mellett kerülhető el a paradoxon, megfontolásai itt is alkalmazhatók. A cikk egyik tétele szerint, ha a kapcsolati mutató érzéketlen a táblázat sormarginálisainak megváltoztatására, akkor a T_{PF} előjele mindig ugyanazt adja, mint az esélyhányados logaritmusa, utóbbi viszont megengedi a Simpson-paradoxont. Rubinék a fundamentális probléma statisztikai megoldását várható értékekkel (külön kezeltre, külön kontrollra) definiálják, ami definíció szerint maga után vonja, hogy a kapcsolati mutató érzéketlen a táblázat sormarginálisainak megváltoztatására, tehát a Rubin-Holland modell eredendően magában hordozza a Simpson-paradoxon lehetőségét. Rudas (2010) felhívja a figyelmet más kapcsolati mutatók használatának lehetőségére, melyek közül néhány nem engedi meg a Simpson-paradoxont. Ha pl. megfigyeléses vizsgálatra épülve populációs szinten szeretnénk a gyógyulásból adódó hasznosságot maximalizálni, akkor nem a hatásátlagokat, hanem a hatásösszegeket kell összevetnünk. Az így definiált kapcsolati mutató a cikk tétele szerint nem engedi meg a Simpson paradoxont.

A fenti tábla segítségével a T hatást definiáló (első) táblát ki tudnánk tölteni, ha ismernénk az alábbi kontrafaktuális adatokat:

	<i>Egészséges</i>	<i>Beteg</i>
Kezelés kontrollba soroltakra	n_{t0}	$n_0 - n_{t0}$
Nem-kezelés kezeltbe soroltakra	n_{c1}	$n_1 - n_{c1}$

5. táblázat A nem megfigyelt, tényellentétes tábla: a hiányzó adatok

A 3. táblázat tehát az ismert 4. táblázat és a nem ismert 5. táblázat összege, így az megfigyelt és meg nem figyelt adatok különválasztásával hiányzó adat problémaként fogalmaztuk meg a

kérdést. A nem megfigyelhető paraméterek száma az elérhető legkisebb, kettő lett (két ismeretlen paraméter a két tényellentétes szituáció miatt mindenképpen adódik).

Ha randomizált kísérletet végzünk, akkor az allokáció független a hatástól, vagyis a gyógyultak aránya kezelés/nemkezelés esetén **várhatóan** ugyanannyi a kezeltbe soroltakra, mint a kontrollba soroltakra, tehát a két utóbbi tábla sorszázalékai **várható értékben** megegyeznek. Ha nem független az allokáció, vagy ha a randomizált kísérlet aktuális allokációját, és az nem ismételt allokációkból várható eredményeit vizsgáljuk, kérdés: mit tudunk mondani a T_{PF} és T közötti eltérésről? Az eltérés dekomponálása (Heckman, Robb, 1985 és pl. friss összefoglalójában Xie 2013) folytonos hatásra ismert, most alkalmazzuk az eljárást bináris hatásra. Ha q jelöli a kontrollba allokáltak (fix!) arányát, akkor

$$\begin{aligned}
 T &= E(Y | do(S=t)) - E(Y | do(S=c)) \stackrel{\substack{\text{teljes várható érték tétellel} \\ A \text{ szerint bontva}}}{=} \\
 &= (1-q)E(Y | do(S=t), A=1) + qE(Y | do(S=t), A=0) - (1-q)E(Y | do(S=c), A=1) - \\
 &\quad - qE(Y | do(S=c), A=0) = (1-q) \times T_{A=1} + q \times T_{A=0} = \\
 &= (1-q) \times \left(\frac{n_{t1}}{n_1} - \frac{n_{c1}}{n_1} \right) + q \times \left(\frac{n_{t0}}{n_0} - \frac{n_{c0}}{n_0} \right) = \left(\frac{n_{t1}}{n_1} - \frac{n_{c0}}{n_0} \right) - \left(\frac{n_{c1}}{n_1} - \frac{n_{c0}}{n_0} \right) - q \left[\left(\frac{n_{t1}}{n_1} - \frac{n_{c1}}{n_1} \right) - \left(\frac{n_{t0}}{n_0} - \frac{n_{c0}}{n_0} \right) \right] = \\
 &= T_{PF} - \text{Torzítás}_{I(\text{baseline különbségből})} - q \times \text{Torzítás}_{II(\text{kezelésre adott reakciók eltéréséből})}.
 \end{aligned} \tag{6}$$

Vagyis a megfigyelt hatás kétféle szelekciós torzítással terhelt jön a tényleges hatásból. Az első típusú torzítás a kezelt és a kontrollcsoportban megfigyelt egészségesarányok eltéréséből jön. Más megfogalmazásban a válasz (Y) és az allokáció (A) kapcsolaterősségét méri a nincs-kezelés ($S=c$) beállítás mellett. Ha $rd(YA)$ -val jelöljük a válasz és az allokáció kockázatkülönbséggel (*risk difference*) mért kapcsolaterősségét, akkor ez a torzítás a

$$\text{Torzítás}_I = rd(YA | do(S=c))$$

formulával írható fel.

Éppen ennek a torzításnak a meglétét definiálta Greenland és Robins (1986, 2009) *noncomparability*-ként, lásd a (4) formulát, és ennek nulla voltával azonosította a *confounding* hiányát. A második torzítás az allokáció és a kezelés eredménye közötti kapcsolatból jön, angol kifejezéssel *treatment effect heterogeneity*, tk. a kezelésre való reagálás eltérése a kezelt- és kontrollcsoport között. Ez is egyfajta szelekciós torzítás, hiszen szintén az allokáció informativitásával függ össze, ám a *baseline* torzítástól függetlenül bármilyen előjelet felvehet. A fenti jelölésekkel ez a torzítás két *risk difference* közötti különbséggé írható fel, ám itt (a tényellentétes alap miatt) ismét nem alkalmazhatók a klasszikus statisztika jelölései, hiszen nem egy együttes eloszlásra vonatkozik a *risk difference* mutató, hanem két különböző lehetséges világ közötti eltérésre. Jelölje ezt a továbbiakban rd' :

$$rd'(YS) = P(Y=1 | do(S=t)) - P(Y=1 | do(S=c)).$$

Most már felírhatjuk a második típusú torzítást:

$$\text{Torzítás}_{II} = rd'(YS | A=1) - rd'(YS | A=0).$$

Tehát az első torzítás egy kapcsolati mutatóként, a második torzítás két kapcsolati mutató értékének A különböző értékei mellett mért változásaként (tk. egy interakciós hatásként, A -nak az YS kapcsolatra gyakorolt hatásaként) fogható fel. Érdemes meggondolni, hogy T és T_{PF} is felírható ezen mutatók segítségével:

$$T = rd'(YS),$$

$$T_{PF} = rd(YS | do(S = t | A = 1) \wedge do(S = c | A = 0)).$$

Az empirikus szociológiai vagy epidemiológiai irodalmak általában az első típusú szelekciós torzítást tárgyalják, és ennek a *baseline* különbségnek a csökkentésére tesznek erőfeszítéseket a kísérlettervezés során is. A Salk-vakcina ismert esetében pl. (Freedman, Pisani, Purves, 2009) a vakcinát a kísérletbe beleegyező szülők gyermekein tervezték kipróbálni, a bele nem egyezők adták volna a kontrollcsoportot. Ám az allokáció így informatív lett volna: a visszautasítók között több volt az alacsony gazdasági-társadalmi státuszú szülő, akik gyermeke viszont a rosszabb körülmények miatt már több kórokozóval találkozott, így szerzett egyfajta alapvédetségét. A *baseline* torzítás itt negatív lett volna – a kezelt, magasabb státuszú gyerekek *baseline* rosszabb esélyekkel indultak a visszautasítóknál – vagyis alulmérték volna a szer hatását. (Meggjegyezhetjük, hogy egy hasonló kísérleti design ma inkább pozitív torzítást eredményezne: éppen a magasabb státuszú szülők között nagyobb a bizalmatlanság az oltásokkal szemben).

A *baseline* torzításra másik példa a *Bursa Hungarica* Felsőoktatási Önkormányzati Ösztöndíjrendszerrel adható. Az ösztöndíj célja az esélyteremtés érdekében a hátrányos helyzetű, szociálisan rászoruló fiatalok támogatása, tehát eleve a rosszabb anyagi helyzetben levő fiatalok kerülnek be a programba. Ha tehát a program hatását szeretnénk kimérni, akkor erre a *baseline* különbségre kontrollálni kell.

A második típusú torzításra, a kezelésre történő eltérő reakcióra jó példa a cégek pályázati támogatásával adható. A nyertes és a nem nyertes/nem pályázó cégek *baseline* vélhetően különböznek. Hiszen a támogatás nélkül is életképesebbek azok, akiket nyertesnek nyilvánítanak, részben azért, mert a pályázaton induló cégek eleve tájékozottabbak, motiváltabbak, részben azért, mert a pályázati elbírálás szempontjai is ezekre a jegyekre épülhetnek. Ugyanakkor a nyertes és a nem nyertes/nem pályázó cégek a támogatásra adott reakciójukat tekintve is különbözhetnek. A nyertes cégek (ismét részben a pályázati elbírálás szempontjai miatt is) épp azok lehetnek, akik egyébként kihasználatlan tartalékkal rendelkeznek (nyersanyag, humánerőforrás), és ezeket a tartalékokat eredményesen ki tudják használni a támogatás birtokában. Vagyis nagyobb hatást ér el náluk a pályázat, mint amit a nem nyertes/nem pályázó cégeken elérne. Fontos implikációja a jelenségnek az, hogy a támogatottak körének kiszélesítése várhatóan csökkentené a támogatás átlagos hatását (!).

Willis és Rosen (1979) cikke is citálható: a szerzők szerint a gimnáziumi/főiskolai képzés hatékonyságának megítélhetősége azért nehéz, mert ezek az oktatási formák eleve azokat a diákokat vonzzák, akik a legtöbbet tudnak profitálni belőle.

A *baseline* torzítást tervezett kísérletesetén sokszor valóban 0-ra tudjuk simítani, mert jobban ismerjük a *response* – a betegség, a cégműködés, az iskolai teljesítmény – meghatározó

faktorait nem-beavatkozás esetén. Ám a *treatment-effect* heterogenitást általában nem tudjuk előzetesen kontrollálni, mert kevésbé ismerjük a beavatkozás (a gyógyszeres kezelés, a cégek anyagi támogatása, az oktatási program) hatásának meghatározó faktorait. Ezzel kapcsolatban Angrist és Krueger (1999) is említi azt a problémát, hogy a bevett, fix-hatás módszerek a kezelés előtt mért kovariánsokra illesztenek, ami csak a baseline heterogenitást tudja csökkenteni.

Randomizáció esetén a kétféle torzítás várható értéke 0, mert a kezelt- és a kontrollcsoport várható válasza megegyezik akkor is, ha mindkettő kezeljük, s akkor is, ha nem:

$$(Y|do(S=t), Y|do(S=c)) \perp A.$$

A T és T_{PF} közötti különbség felbontását leíró (6)-os egyenletben a q szerepe elsőre talán váratlannak tűnik: a félremérés értékét befolyásolja a q , vagyis az allokációs arány. Ám fontos megjegyezni, hogy itt az allokációt, tehát a q -t is fixnek tekintjük. Tehát nem érvényes pl. az a meggondolás, hogy a második típusú torzítás hatását minimalizálni tudjuk, ha q -t, a kontrollcsoport nagyságát minimalizáljuk. A q itt fix, a T maga is ennek függvénye, és mi erre a konkrét (vagy várható, ha ismétlődő mintavételt teszünk fel) mintára és konkrét T -re végzünk megfigyelést. Vagyis a látszólagos hatást nem, de a valódi hatást befolyásolja az allokációs arány. Ez a tulajdonság a Rubin-Holland modellből jön, ahol a T tulajdonképpen a kezelt és a kontroll csoportra vett hatás *csoportnagyságokkal súlyozott* átlaga. Tekintsük a kezelt- és a kontrollcsoportot két egyedtypusként (pl. az első típusba tartozó alany, ha rábízzuk a döntést, az első kezelést preferálja, a másik típus a második kezelést). Ekkor, feltéve, hogy a két csoport releváns módon eltér az *outcome*-ot illetően (akár *baseline*, akár a kezelésre adott reakciót tekintve), a T és T_{PF} közötti eltérés kézenfekvő módon függ attól, hogy milyen a két típus aránya a mintában.

Ha a két csoport között nincs releváns eltérés (a két torzítás mindegyike 0), akkor a q értéke irreleváns, a megfigyelt hatás a ténylegessel egyezik meg.

Egy konkrét számpéldára, a veseköves-vizsgálatra térve: tegyük fel, hogy a kezelt- és kontrollcsoport eltérése kizárólag a nagy/kis vesekövek megoszlásából adódik, vagyis hogy a méretre feltéve az allokáció már nem informatív (lásd a causality.xls torzítások c. munkalapját). Ezt a (3) formulával definiált összefüggést neveztük korábban erős ignorálhatóságnak: a kőméretre feltéve az Y hatás független az allokációtól. E feltevés mellett a tényellentétes adatok is kiszámolhatóak, hiszen pl. a PCNL-lel műtöttek (kezelték) nyílt műtét esetén kapott gyógyulási arányát ki tudjuk számolni abból, hogy ismerjük a kőméret szerinti összetételüket, és láttuk, hogy a kis illetve nagy kővel rendelkezők hogyan reagáltak a nyílt műtétre. A számításokat lásd az excel file-ban. Az eredmények szerint

$$T_{PF} = 0,046$$

$$\text{Torzítás}_I = 0,11$$

$$\text{Torzítás}_{II} = -0,01$$

$$T = T_{PF} - \text{Torzítás}_I - q \times \text{Torzítás}_{II} = 0,046 - 0,11 - 0,5 \times (-0,01) = -0,054$$

Vagyis bár e megfigyelt hatás pozitív, vagyis eszerint a PCNL preferálandó, a valódi hatás a nyílt műtetet mutatja előnyösebbnek. A PCNL hatásának e felülmérése két torzítás eredőjeként adódik. A *baseline* különbség a két csoport között nagy, és pozitív előjelű, vagyis az első torzítás a hatás felülmérése irányába hat (mert a kezelték között jobb kilátású - kisebb köves - betegek vannak). A kezelésre adott reakció két csoport közötti eltérése (második torzítás) elhanyagolható, egyébként negatív előjelű, tehát a hatás alulbecslésének irányába mutat. Ez a negatív előjel arra utal, hogy a PCNL hátránya a nyílt műtéthez képest az aktuálisan PCNL-lel műtötteken (kezeltékbe soroltakra) nagyobb, mint a nyílt műtéten átesetteken (a kontrollcsoporton).

Greenland és Robins fent idézett munkái (1987, 2009) csak az első típusú torzítás jelenlétére vezetnek vissza a *non-comparability*, vagyis a *confounding* jelenlétét. Pedig a második típusú szelekciós torzítás is a két csoport eltérő összetételéből, össze nem vethetőségéből ered, és ez is okozhatja a T és T_{PF} közötti eltérést, akár előjel-eltérést. Hiszen mindkét torzítás előjele lehet pozitív és negatív is, és egymástól függetlenül különböző irányokba torzíthatják a hatásmutatót. Az 1. táblázatbeli adatokkal és a kőméretre vett erős ignorálhatóság feltevésével láttuk, hogy az első torzítás értéke 0,11, a másik -0,01. Az erős ignorálhatóság feltevése itt azt implikálta (lásd a *causality.xlsx* torzítások munkalapját), hogy a tényellentétes tábla az alábbi módon alakult:

	<i>Gyógyult</i>	<i>Nem gyógyult</i>
Ha nyílt műtéttel operáltam volna PCNL-lel műtötteket	310	350-310
Ha PCNL-lel operáltam volna nyílt műtéten átesetteket	256	350-256

6. táblázat A tényellentétes tábla, ha a kőméretre feltéve fennáll az erős ignorálhatóság

Ha ez a tényellentétes tábla másképp alakul, akkor a két torzítás is változik (lásd a *causality.xlsx* hiányzó_adat munkalapján a változtatható input-cellákat). Ha pl. a 310 helyébe 40-et írok (vagyis radikálisan lerontom a kezelt csoport nem-kezelés esetén adódó gyógyulását), akkor a torzítások előjele megfordul. Az első torzítás értéke -0,67 lesz (a *baseline* torzítás erős, és előjelénél fogva a hatás alulbecslése irányába hat, mert a kezeltékbe súlyosabb esetek kerültek), a másik torzítás 0,76 lesz (az erős torzítás a hatás felülbecslése irányába hat, mert a kezelték közé soroltak erősebben reagálnak a kezelésre).

A probléma pontos paraméterezésével és a két hatás eltérésének felbontásával tehát kézenfekvő módon, könnyen értelmezhető keretben tárgyalható a megfigyeléses vizsgálatokban jelentkező hatástorzítás és annak tényezői. Ha a *confounding*-ot *noncomparability* értelemben használjuk, akkor az ekvivalensnek tekinthető a két torzítás közül legalább az egyiknek a jelenlétével. A *noncollapsibility* viszont közvetlenül nem tárgyalható ezen a kereten belül, nincs harmadik, *confounder* változónk, továbbá a T és T_{PF} eltéréséből nem következtethetünk bármely konkrét változó szerinti rétegzett elemzés eredményére. Viszont az erős ignorálhatóság feltevésével, ahogy a veseköves példán láttuk, ez a keret felhasználhatóvá válik *noncollapsibility* -problémára is.

Vegyük észre, hogy T és T_{PF} megegyezése nem jelenti a *comparability* teljesülését: a torzítások eltérő előjellel kiolthatják egymást, így T és T_{PF} azonossá válhat, annak ellenére, hogy a *confounding* fennáll. A fenti példa folytatásaként: ha az 6. táblázatban a 310 helyébe

240-et írunk, akkor az Excel-file szerint a két torzítás kioltja egymást, és a megfigyelt hatás (mintegy véletlenül) éppen a valódi hatással egyezik meg.

Érdemes megjegyezni, hogy bár a T és T_{PF} közötti különbség felbontását leíró fenti formula az ismertebb, a felbontás másképp is átalakítható, szintén jól interpretálható tagokkal (Németh, Rudas, 2014):

$$T = T_{PF} - (1-q) \times \left(\frac{n_{c1}}{n_1} - \frac{n_{c0}}{n_0} \right) - q \times \left(\frac{n_{t1}}{n_1} - \frac{n_{t0}}{n_0} \right).$$

Itt az első torzító tényező szintén a *baseline* különbség, a második tényező viszont nem a kezelés kiváltotta gyógyultarány-változásban jelentkező eltérés, hanem a kezelés esetén megfigyelt gyógyult-arányok eltérése. Ennek az eltérésnek a 0 voltát követelte meg a *baseline* azonosság mellé Greenland és Robins a teljes összevethetőség teljesüléséhez, lásd az (5) formulát. A q szorzótényezők kézenfekvő voltának indoklásához tegyük fel, hogy nincs *baseline* torzítás, de a második típusú torzítás fennáll. Ekkor a valódi hatáshoz képest a T_{PF} , ami a kezelés megítélésénél csak a kezeltbe allokáltakat veszi figyelembe, annál nagyobb torzítást mutat, minél nagyobb az általa figyelembe nem vett csoport (a kontroll aránya, q). Ugyanis a valódi hatás tulajdonképpen súlyozott átlaga a rétegenkénti hatásnak, a T_{PF} viszont nincs tekintettel a rétegek arányára, így ha a kontroll nagy részét adja a mintának, akkor a T_{PF} helytelenül vesz sok információt a kezeltéken mérhető treatment-hatásról. És fordítva: ha inkább a kezelték adják a mintát, akkor nem jelent nagy torzítást, hogy a T_{PF} csak a kezeltéken mér hatást.

Az $(1-q)$ szorzótényező hasonlóan indokolható: ha a kezelés nélküli hatás nem egyforma, akkor a T_{PF} , amely csak a kontrollba allokáltakat veszi figyelembe a kezelés nélküli helyzet megítéléséhez, annál nagyobb hibát okoz, minél nagyobb az általa figyelembe nem vett csoport (a kezelték) aránya.

Végül röviden kitérnék annak a szemléletnek az előnyére, ami a T és T_{PF} eltérését hiányzó adat problémaként definiálja, lásd a tényleges hatást definiáló 3. táblázat felbontását a megfigyelt 4. és meg nem figyelt 5. táblázatra. Ez a szemlélet lehetőséget ad a megfigyelt hatás megítélésére. Pl. adott megfigyelt hatás mellett feltehető az a kérdés, hogy ha a hatás valójában nem áll fenn (nincs hatása a kezelésnek), akkor milyen értékeket tartalmazhat a meg nem figyelt táblázat? A meg nem figyelt táblázat többféleképpen is kitölthető, hiszen két ismeretlen értéket tartalmaz (két paraméterrel írható le). A causality.xls file hiányzó_adat c. munkalapja a veseköves példán szemlélteti ezt: a nem megfigyelt táblázat celláit mi tölthetjük ki, s ennek függvényében fog változni a kétféle torzítás, így a T valódi hatás is.

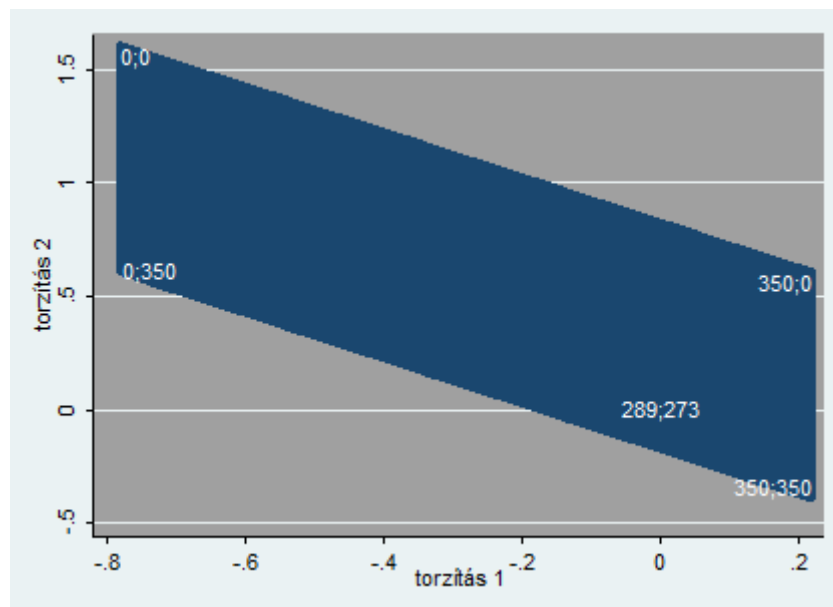
Ennek megfelelően a T és T_{PF} közötti eltérés is két paraméterrel, pl. a kétféle torzítással adható meg. Ha viszont feltesszük, hogy valójában nincs hatás, és a megfigyelt hatás csak a torzítások eredményeként jött létre, akkor egyetlen paraméter változtatható már csak. Ilyen és ehhez hasonló kérdések könnyen válaszolhatók meg ebben a keretben: ha az első torzítás értékét rögzítjük (pl. a design megfelelő tervezésével, háttérinformációkat felhasználva a torzítást 0 közeli értékre tolva), akkor mi lesz a második torzítás minimális értéke, feltéve, hogy valójában nincs hatás. Ennek a torzításnak, illetve a neki megfelelő nem-

megfigyelt tábla “hihetőségének” a megítélésével dönthetünk aztán a nincs-hatás hipotézis elfogadhatóságáról. Egy példát mutatok a causality.xls hiányzó_adat_II c. munkalapján, a második típusú torzítás minimalizálásának (egy tulajdonképpen lineáris optimalizálási problémának) a megoldására az Excel Solver bővítményét alkalmazva. A veseköves kezelés esetében azt látjuk, hogy ha feltesszük, hogy az általunk látott hatás (a kezelés, a PCNL pozitív hatására vonatkozóan) csak a torzítások eredményeként jött létre, és valójában nincs hatás, és ha arról is meg vagyunk győződve, hogy sikerült a *baseline* különbséget a kezelt és kontroll csoport között kisimítani, akkor a második torzítás értéke minimálisan 9% kell, hogy legyen. Vagyis a kezelték ennyivel jobban reagálnak a kezelésre mint a kontrollok, eltérő összetételüknek fogva. Ez elég nagy értéknek tűnik, a megfigyelt 4,5%-os hatáshoz viszonyítva. A meg nem figyelt táblára vonatkozóan azt látjuk, hogy az első típusú torzítás 0 volta annak felel meg, hogy a kísérletben PCNL-lel műtötteket nyílt műtétnak alávétve a kísérletben is nyíltan műtötteknek teljesen megfelelő gyógyulási arányt kapunk. A második típusú torzítás pedig abban nyilvánul meg, hogy a nyílt műtéttel műtötteket PCNL-lel műtve a ténylegesen PCNL-lel műtötteknél jóval rosszabb eredményt kellene kapni (83% helyett 73%-os gyógyulási arányt).

Említettem, hogy a T és T_{PF} közötti eltérés két paraméterrel írható le, ennek egyik lehetséges módja a kétféle torzítás, mint paraméter használata. A vizsgálatokat vagy vizsgálat-típusokat, designokat (utóbbiak esetében várható torzításokat megadva) egy, a két torzítás által meghatározott síkon helyezhetjük el (Németh, Rudas, 2014). A randomizált kísérlet, mint típus pl. az origóban van (várhatóan mindkét szempontból torzítatlan), de egy-egy konkrét randomizált kísérlet hordozhat torzítást, tehát az origótól távolabb is elhelyezkedhet, nyilván a mintanagyságtól függő távolságban. Ha a *baseline* torzítást jól tudjuk kontrollálni, akkor a vizsgálatoknak megfelelő pontok első koordinátája 0 közelében lesz, a függőleges tengely közelében helyezkedve el. Megfigyeléses vizsgálatnál mindkét torzítás lehet pozitív és negatív is. A $Torzi\acute{t}\acute{a}s_I$ értéke elméletileg -1 és 1 között változhat, a $Torzi\acute{t}\acute{a}s_{II}$ értéke -2 és 2 között. A kétféle torzítás előjele egymástól függetlenül változhat, de abszolút értékük nem, így a kapott síkpontok nem az értékkészletek Descartes-szorzatán (nem a 4x2-es téglalapon) helyezkednek el.

A 4. ábra a veseköves betegek kezelésének példáján illusztrálja ezt. Az ábra a kétféle torzítás lehetséges értékeit mutatja a tényellentétes tábla különböző kitöltési mellett (ahogyan a causality.xlsx hiányzó_adat munkalapján kiszámolhatjuk). Itt sem téglalapot, hanem paralelogrammát kapunk. Látható, hogy az első torzítás értéke -0,78 és 0,22 között lehet, a második torzításé -0,39 és 1,6 között. A paralelogramma öt pontját emeltem ki: a torzításmentes szituációt (random allokáció esetén várhatóan ide esik a vizsgálat), és a maximális/minimális torzítások mellett adódó négy sarokpontot. Az öt ponthoz megadtam, hogy a tényellentétes tábla milyen kitöltése mellett adódnak; az első szám a kezeltbe soroltak közül a gyógyultak száma, ha nem kezeltük volna őket, a másik szám a nem kezeltbe soroltak között a gyógyultak száma, ha mégis kezeltük volna őket. A megfigyelt hatás itt 0,046 volt, a valódi hatás a tényellentétes adatokat változtatva -0,48 és 0,52 között mozog. A minimumértéket (a T és a kétféle torzítás közötti lineáris függvénykapcsolatból adódóan) az első torzítás maximális értéke mellett maximalizált második torzítás esetén veszi fel

(paralelogramma jobb felső sarka), a maximumértéket a minimális első torzítás mellett minimalizált második torzítás esetén (paralelogramma bal alsó sarka).



4. ábra A kétféle torzítás lehetséges értékei a tényellentétes tábla különböző kitöltései mellett, veseköves példa

Fent a hatás megítélésének egy módjaként a nincs-hatás hipotetikus szituációban kapott jellemzők „hihetőségét” említettem. A hihetőség kézenfekvő módon egy harmadik, potenciális *confounder* változó bevonásával is számszerűsíthető: ha a valódi tábla független (nincs hatás), milyen erősnek kell lenni egy *confounder* hatásának az allokáció-válasz kapcsolatra, ha nincs kezelés egyáltalán (*Torzítás_I*) illetve milyen erősnek kell lennie a *confounder* hatásának a kezelés-válasz kapcsolatnak az allokáció egyes rétegei között megfigyelt eltérésére (*Torzítás_{II}*)? Hihető-e ez a két hatáserősség? Ezzel kapcsolatban két szakirodalmi forrás említhető. Az egyik Cornfield és társainak klasszikus, 1959-es cikke. A cikk tulajdonképpen a „Korreláció nem implikál okságot” törvény mellé nyújt konkrétabb fogódzót azzal, hogy szükséges feltételeket fogalmaz meg a megfigyelt kapcsolat előjelének egy harmadik változó bevonását követő megfordulására (a Simpson-paradoxonra). A szerzők tétele szerint ha egy *A* változó (dohányzás) valójában nem hat az *E* változóra (tüdőrák), akkor a tényleges ok (*C*) valószínűségének a következő feltételt kell kielégíteni:

$$P(C|A)/P(C|nem A) > P(E|A)/P(E|nem A).$$

Vagyis ha a dohányzók rák-kockázata 9-szerese a nem-dohányzókénak, akkor Fisher génje²⁸ előfordulásának a dohányzók között több mint 9-szer gyakoribbnak kell lenni, mint a nem dohányzók között. A causality.xls file első munkalapját tekintve a vesekövek kezeléséről: a nyílt műtéttel operáltak relatív kockázata PCNL-esekhez képest 94%, legalább ilyen erősnek

²⁸ Utalás Fisher azon érvére (lásd pl. Freedman et al, 2005), hogy nem a dohányzás okozza a tüdőrákot, hanem egy olyan gén, ami a dohányzásra való rázokást is elősegíti.

kell lenni a műtéttípus kapcsolatának a kórérettel is, hogy az előjel forduljon. És valóban, a relatív kockázat az utóbbi esetben 30%.

Milo Schield (1999) a feltétel egy jobban használható változatát adja: a hatás kioltásához vagy megfordításához szükséges feltétel, hogy a „valódi ok” hatásnagysága (ismét kockázat-különbséggel mérve) legalább a mért legyen:

$$P(E|C) - P(E|nem C) \geq P(E|A) - P(E|nem A)$$

Schild feltétele is áll a veseköves példán: a sikertelen műtét kockázatának különbsége nyílt és PCNL-műtötteket összevetve 4.5%, míg nagy és kis kőnagyság összevetésében 16%.

Rosenbaum és Rubin (83) ennél többet is tud mondani. Egy eljárást ajánlanak annak megítélésére, milyen érzékeny lehet a megfigyelt bináris hatás egy meg nem figyelt kategoriális *confounder*-re (miközben megfigyelt *confounderek*re is kontrollálnak). A meg nem figyelt *confounder*-ről felteszik, hogy rá is feltéve az allokáció hatása már erősen ignorálható. A valódi hatás *confounder*-től való függését parametrizálják két paraméterrel (mennyire növeli a *confounder* a kezeltbe allokalás esélyét, ill. mennyire különbözik a pozitív *outcome* esélye a *confounder* két értéke mellett, külön kezeltre és külön kontrollcsoportra). A paraméterbeállítások rögzített (pl. hihetőnek tartott) értékei mellett adnak aztán maximum likelihood becslést a valódi hatásra. Konkrét orvosi alkalmazási példájukból az derül ki, hogy a hatásirány megfordításához (az allokációval és az *outcome*-mal mért kapcsolatra vonatkozóan egyaránt) irreálisan erős hatású *confounder*-t kellene feltenni, vagyis a konklúzió szerint nem vélelmezhető, hogy a valódi hatás előjele eltér a megfigyelttől. Úgy is fogalmazhatnánk, hogy a szerzők annak vizsgálatára adnak eljárást, milyen megbízhatósággal lehet az adott megfigyeléses vizsgálat a Simpson-paradoxon alege. A causality.xls RR_simpson munkalapján a vesekő-kezelés példáját paramétereztem fel Rosenbaum és Rubin paramétereivel. Az adott paraméterértékek az ismert *confounder*, a vesekő mérete mellett kapott értékek, de változtatásukkal más szituációk is megítélhetők. A munkalapon a valódi *T* hatás értéke és a *Torzítás_I*, *Torzítás_{II}* értéke is megjelenik, tehát a fenti szemléletbe is beilleszthetők az eredmények. Vagyis ismét azt látjuk, hogy gyümölcsözően egyesíthetők a *confounding*-probléma egyes megközelítései a Rubin-Holland modellel.

5. Grafikus modellek és kauzalitás

Az előző fejezetben a *confounding*-nak az össze-nem-vonhatósággal (*noncollapsibility*-vel) adott definíciója kapcsán érintettem azt a kérdést, hogy vajon mindig kontrollálni kell-e azokra a változókra, amikre nézve nem tehető meg az összevonás, vagyis amikre vett feltételes rétegekben a kapott asszociációs mérőszámok eltérnek a nyers, nem rétegzett mérőszámtól. Másképp: a kontrollált vagy a nyers hatásmutató megbízható-e? Ennek az irodalomban *adjustment problem*-ként ismert problémának Judea Pearl (2009) által adott megoldásakor az asszociációk mérésén felül a hatásirányokat is definiálnunk kell. Ezek az irányított kapcsolatok grafikus modellekkel szemléltethetők jól, ahol a gráf csúcsai az egyes változókat, a gráf élei pedig a változók közötti kapcsolatokat reprezentálják. A gráf két csúcsa közötti hiányzó él a két változó feltételes függetlenségéeként értelmezhető, a feltételes függetlenség feltételhalmaza (és az összes fennálló, nem feltétlenül páronkénti feltételes

függetlenség) a gráfból ún. Markov-tulajdonságok segítségével olvashatók ki. Az alábbiakban először röviden a grafikus modellekkel kapcsolatos saját munkámat mutatom be, mintegy beágyazva azt az okság problematikájába.

A kategoriális változók közötti, feltételes függetlenségekkel leírható kapcsolatokat megragadni képes grafikus modelleket vizsgáltam már disszertációmban is. E grafikus modellek oksági láncok kezelésére, direkt és indirekt hatások megkülönböztetésére alkalmasak. Könnyen vizualizálható és jól interpretálható módon jelenítenek meg komplex rendszereket, és modulárisak, vagyis e komplex rendszereket egyszerűbb részek összekapcsolásával építik fel. Az utóbbi években született publikációink módszertani alapkérdése a kontingenciatábla olyan paraméterekkel történő felparaméterezésének megtalálása, amelyekkel könnyű e grafikus modellek megadása. A felparaméterezés a táblát egy-egyértelműen meghatározó mutatók megadását jelenti. Ez a kérdés megkerülhetetlen, ugyanis (1) a grafikus modell ezeknek a paramétereknek bizonyos megszorításaival adható meg. Továbbá (2) azokban a gyakori esetekben, amikor a cél egy jól illeszkedő gráf (azaz az ezáltal prezentált modell) élei (a modellben szereplő hatások) elhagyhatóságának a tesztelése, akkor ezt a paramétermegszorítások fokozatos feloldásával (*stepwise* - lépegetve) érhetjük el. Ugyanennek az esetnek a megfelelője, amikor rosszul illeszkedő gráf javítása a cél, és bizonyos élek behúzásának hatását teszteljük a paramétermegszorítások fokozatos feloldásával. Végül (3) a modellhez tartozó eloszlás értelmezése a becsült paraméterek interpretációjával történik. Mindhárom esetben azon múlik az eredményünk, hogy a tábla paraméterezése milyen paramétereket tartalmaz. Mint szerzőtársaimmal a Biometrikában 2010-ben megjelent publikációnkban (Rudas, Bergsma, Németh, 2010) megmutattuk, ezek a problémák a grafikus modellek bizonyos osztályai esetén marginális loglineáris paraméterek megadásával jól kezelhetők. A marginális loglineáris paraméterek a hagyományos loglineáris paraméterektől abban térnek el, hogy nem csak az összes változó együttes eloszlása, hanem bármely marginális eloszlás mellett definiálhatunk paramétereket, így finomabban tudunk hatásokat modellezni velük. A cikkben megmutattuk, hogy az eredmények alkalmazhatók az irányított és irányítatlan hatásokat egyaránt kezelni képes láncgráf-modellek egy ismert típusára is.

A Bulletin of Sociological Methodology-ban 2013-ban elsőszerzőségemmel megjelent cikkben (Németh, Rudas, 2013b) a grafikus modellek mobilitáskutatási felhasználhatóságát mutattuk be. A hangsúly itt az alkalmazási kontextusban való elhelyezésen volt, tárgyaltuk a mobilitáskutatás módszertanának történeti változásait, a tartalmi problémák és az alkalmazott eszközök megfigyelhető koevolúcióját is. Saját alkalmazásként Magyarország, az USA és Csehszlovákia 1992-es ISSP adatain vizsgáltuk a státuszelérés folyamatát. Elemzésünk egybecseng azokkal a korábbi eredményekkel (pl. Szelényi Szonja), amelyek sikertelennek ítélték a szocialista rezsimnek a társadalmi egyenlőtlenségek intergenerációs átörökítésének csökkentésére tett erőfeszítéseit.

A kategoriális grafikus modellekről első szerzőségemmel készült összefoglaló munkánk a *Sociological Methodology* marginális modelleket tárgyaló tematikus különszámában jelent meg 2013-ban, felkért szakértők hozzászólásaival és a rájuk adott viszontválaszunkkal (Németh, Rudas, 2013a). A tanulmány fő célja a módszer bemutatása volt a marginális

modellezés nyújtotta egységes statisztikai kereten belül. A marginális modellek közé tartozó kategoriális útmodellek egy példajaként vizsgáltuk a státuszelérési folyamatot, a liberális és a konzervatív jóléti rendszerek összevetésében. A cikk kísérőjeként létrehoztam egy honlapot is (<http://nemethr.web.elte.hu/discrete-graphical-models/>), ahol a cikk alkalmazási példái találhatóak, adatbázisokkal és R-scriptekkel együtt.

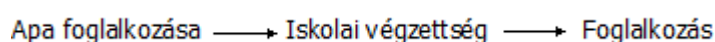
A fenti cikkekben röviden érintettük a modellek kauzális interpretációjának problémáját. Megmutattuk pl., hogy irányított gráfok esetén az élek irányítása bizonyos speciális esetekben megítélhető az adatokból (különböző modellek illeszkedését összevetve), de általánosságban, az adatok nem elegendőek az irányítás eldöntésére. Ide tartozik a Markov-ekvivalens gráfok problémája: ezek különböző irányítású éleket tartalmazó, mégis, a Markov-tulajdonságok által ugyanazokat a feltételes függetlenségeket implikáló gráfok.

Fókuszunkban ugyanakkor nem a modellek kauzális értelmezése, hanem a kategoriális változókra felírt modellek paraméterezhetősége állt. Az általunk követett iskola (lásd pl. Whittaker, 1990, Lauritzen, 1996, Cox, Wermuth, 1996, Andersson, Madigan, Perlman, 2001, Richardson, 2003 és Drton, 2009) a gráfokat egészként, a hiányzó élek által implikált feltételes függetlenségekkel értelmezi, szándékosan elkerülve az egyes irányított élek inherens kauzális interpretációjának problémáját. Ez azt jelenti, hogy egy adott élnek nincs önálló jelentése, az csak a többi él meglétének/hiányának ismeretében interpretálható. A grafikus modellek ettől eltérő, kauzális tárgyalása elsősorban Judea Pearl (elsősorban 2009) nevéhez fűződik. A megközelítés sajátosságát jelzi, hogy Pearl nem grafikus modelleknek, hanem expliciten oksági gráfoknak (*causal graphs*) nevezi a modelleket. Pearl oksági gráfjai segítségével detektálhatóvá tudjuk pl. tenni a *confounding*, a szelekciós torzítás vagy a mediáció (közvetett hatáslánc egy közbeiktatott változón keresztül) jelenségét, és választ adhatunk arra a fejezet elején említett kérdésre, hogy kontrolláljunk-e bizonyos változókra. Ugyanakkor, mint látni fogjuk, ezek a lehetőségek akkor állnak elő, ha az oksági gráf megkonstruálható – a konstrukció viszont a világról alkotott alapos háttértudásunkat igényli, az együttes eloszlás megfigyelésénél is többet.

Pearl (2009) kritizálja a fent említett, a grafikus modelleket az általuk implikált feltételes függetlenségi állításokkal definiáló, általa uralkodónak nevezett „asszociációs” paradigmát. Szerinte közvetlenül a világról való tudásunkra alapozott oksági kapcsolatok megfogása a lényeges, a feltételes függetlenségek csak ezek következményei. Ebben a megközelítésben a gráf nyilai tényleges mechanizmusokat reprezentálnak, ahol az „ok” manipulálásával (itt az okság manipulációs fogalma jelenik meg!) az „okozat” elkerülhetetlenül változik. Ezek az információk az együttes eloszlás, az asszociációk megfigyeléséből nem nyerhetők ki, annál többet mondanak, hiszen tényellentétes helyzeteke írhatnak le (lásd még az előző fejezetben Pearl do-operátorának bevezetésének indoklását). Pearl oksági gráfjai tehát az asszociációs paradigma grafikus modelljeinél több megszorítást tartalmaznak. Ugyanazok a Markov-tulajdonságok használhatók a feltételes függetlenségek kiolvasására (Pearl, 2009, pp. 24.), de az oksági mechanizmus jóldefiniáltságához szükséges tulajdonságokat is megköveteljük: a nyilak direkt oksági hatásként interpretálhatók kell, hogy legyenek (ezt tényellentétes megközelítésben definiálják), és bármely változópár közös okának szerepelnie kell a gráfon

(Hernán et al., 2004). Ez utóbbiak nyilvánvalóan nagyon erős feltételek, sok háttértudást követelnek, nem elég megítélésükhöz néhány változó együttes eloszlásának megfigyelése.

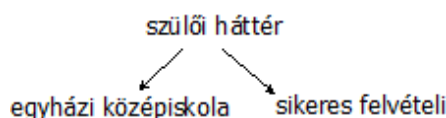
Most rátérhetünk kiinduló problémánkra (*adjustment problem*): annak kérdésére, hogy oksági hatás mérésekor mikor érdemes kontroll alatt tartani azokat a változókat, amikre kontrollálva egyébként megváltozik a nyers asszociáció értéke (*non-collapsibility*). Az ok, okozat és a harmadik, potenciális kontrollváltozó oksági viszonyával kapcsolatban három logikai alapváltozat képzelhető el (Pearl, 2009, Hernan et al, 2004), lásd 7., 5., 6. ábra (ennél komplexebb struktúrák is kezelhetők ebben a keretben, hosszabb láncokkal stb., de ezekre az alapesetekre visszavezethetők). Az ok és az okozat közötti direkt kapcsolatot jelző nyilat mindhárom esetben kihagytam (így erősebbek a példák), de azok szerepelhetnek is. E logikai viszony alapján dönthető el, hogy a harmadik változóra érdemes-e kontrollálni.



5. ábra Közvetítés (*mediation*)

Közvetítés esetén (lásd 5. ábra) a harmadik változó (iskolázottság) az ok (apa foglalkozása) és okozat (utód foglalkozása) közé ékelődik az oksági láncban. A kontrollált hatás a direkt hatást méri csak (ez az ábrán adott modell esetén a hiányzó harmadik nyíl miatt zéró), a nyers hatásba az indirekt hatás is belejátszik. Mindkét hatás jól interpretálható, és más-más kutatási kérdésekre ad választ. Az indirekt hatás az oktatáspolitikát érdekelheti, ha az esélyegyenlőség biztosításával a származás szerepét kívánja csökkenteni. A direkt hatás a társadalmi tőkének az oktatási rendszert megkerülő csatornákon történő átöröklésével, pl. az állásszerzés diszkriminatív jellegével (számítanak-e családi informális kapcsolatok) is kapcsolatban lehet.

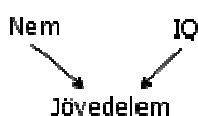
A következő ábra (6. ábra) a pearl-i értelemben vett *confounding* igazi esete: itt a harmadik (kontroll) változó hat mind az okra, mind az okozatra, ezért látszólagos kapcsolatot hoz köztük létre. Példánkban az egyházi középiskolák teljesítményét kívánjuk mérni, de a szülői háttérben jelentkező önszelekció (ha a magasabb státuszú szülők preferálják inkább ezt az iskolatípust) a hatás felülméréséhez vezet. Az ábrán az egyházi iskola és a sikeres felvételi között nincsen nyíl, tehát bár van marginális kapcsolat, a direkt kapcsolat teljesen eltűnik a szülői háttérre kontrollálva. (Az oksági irodalomban gyakran idézett példa még erősebb: a dohányzás megsárgult ujjakat és tüdőrákot is eredményez, ezzel a sárga ujj és a tüdőrák között látszólagos kapcsolatot hozva létre.) A három logikai variáció közül ez az Pearl szerint, ami mindenképpen kontrollálást kíván.



6. ábra *Confounding*

Ezzel szemben az utolsó variáció esetén éppen a kontrollálás vezet be ál-összefüggést, és a nyers kapcsolat méri a „valós” kapcsolatot. A 7. ábra példája jól mutatja ezt: tegyük fel hogy

a nem (férfiak előnyben) és az IQ (magasabb IQ-júak előnyben) is hatással vannak a jövedelemre. A nem és az IQ között nincsen nyers kapcsolat, de a jövedelmet rögzítve kapcsolat jelenik meg. Pl. ha magas jövedelmi kategóriában vagyunk, és tudjuk, hogy nőről van szó, akkor megnő a magas IQ valószínűsége. Ezt a logikai viszonyt az élek irányítására utalva Pearl ütközésnek (*collider*) nevezi. Az irodalomban fellelhető gyakori példa ugyanerre: tudjuk, hogy a nagy sebesség és az alkoholfogyasztás is autóbalesetet okozhat. E két változó között nincs nyers összefüggés, de ha tudjuk, hogy autóbaleset történt, akkor függővé válnak: az egyik cáfolata maga után vonja a másik nagyobb valószínűségét. Ugyanez igaz, ha nem az autóbalesetre, hanem az az által direkt módon befolyásolt más változóra kontrollálunk. Ütközés esetén tehát nem célszerű kontrollálni, mert a feltétellel „mesterséges” kapcsolatot teremtünk.



7. ábra Ütközés (*collider*)

Ezeknek az építőköveknek a segítségével Pearl (2009) komplex gráfok esetére is érvényes algoritmust definiál annak eldöntésére, mely változókra kell kontrollálni. Az alapötlet az, hogy a kontrollálás segítségével minden nem-oksági és közvetett utat (*confounder, mediator*) le kell zárni, és minden érdeklődésre számot tartó oksági utat (*collider*) nyitvahagyni. Ismét hangsúlyoznám, hogy mindez az oksági gráf megkontsruálhatósága, tehát lényeges háttértudás megléte mellett alkalmazható.

Végül a *collider* esetét vegyük részletesebben szemügyre. Az irodalom *admission-*, *selection-collider stratification-* vagy Berkson-féle torzításként ismeri a problémát. A fenti két példában (jövedelem alakulása, autóbaleset) két (korrelálatlan) VAGY-gyal kapcsolt kritérium alapján kerültek „kiválasztásra” (magas jövedelem elérése vagy autóbaleset megtörténte) az alanyok. Ilyenkor a kiválasztottak között a két kritérium negatívan korrelál (ahogy fent láttuk: ha megtörtént autóbaleset, és tudom, hogy nem hajtott gyorsan a sofőr, akkor megnő alkoholfogyasztás valószínűsége). Berkson (1946) ismert problémája is ide tartozik: ha két, a populációban korrelálatlan (B1 és B2) betegség egyaránt kórházi felvételt indokol, akkor a hospitalizált betegek között negatív korrelációt tapasztalunk a két betegség között. Ezért pl. érvénytelen eredményekre vezetnek azok az eset-kontroll vizsgálatok, ahol a B1 betegségben szenvedő kórházban kezelt betegek, mint esetek mellé B2-ben szenvedő betegeket választunk ugyanabból a kórházból kontrollnak: B1 rizikófaktor a B2-vel is (negatív) összefüggést fog mutatni, akkor is, ha a populációban nincs köztük kapcsolat.

Ha a beválogatási kritériumok ÉS-sel kapcsoltak, akkor a populációban korrelálatlan kritériumok a beválasztottak között pozitívan fognak korrelálni. Biológiai példával (Cole et al, 2010): ha egy környezeti körülmény és egy tőle független, random elosztott gén növeli egy betegség kockázatát, még hozzá erős pozitív interakcióban, akkor a feltételes hatás pozitív: a betegek között a gén megléte valószínűsíti a környezeti körülmény fennállását is.

A VAGY- és az ÉS-kapcsolás formalizálható, és egyben általánosítható is. Az alábbiakban bináris változók esetén (Németh, Rudas, 2014 alapján) a feltételes hatás előjelét adjuk meg a bevásárlási valószínűségek függvényében.

Legyen az alábbi tábla a Nem és az IQ populációban megfigyelt együttes eloszlása:

	Magas IQ	Alacsony IQ
Férfi	a	b
Nő	c	d

7. táblázat A marginális tábla

A két változó kapcsolatát mérhetjük (marginális) esélyhányadossal (MEH), ennek értéke

$$MEH = \frac{ad}{bc}. \quad (7)$$

A marginális kapcsolat előjele ennek értékétől függ, MEH=1 esetén független a két változó, 1-nél nagyobb MEH esetén a férfiaknak, 1-nél kisebb MEH esetén a nőknek nagyobb az esélye a magas intelligenciahányadosra. Tegyük most fel, hogy a MEH értéke 1-

Legyenek a magas jövedelemhez vezető cellánkénti „bevásárlási” arányok p, q, r és s (0 és 1 közötti számok), vagyis a magas jövedelem esetén megfigyelt feltételes tábla:

	Magas IQ	Alacsony IQ
Férfi	p*a	q*b
Nő	r*c	s*d

8. táblázat A feltételes tábla

A feltételes esélyhányados (FEH) értéke

$$FEH_{jöv=magas} = \frac{pasd}{qbrc} = MEH \times \frac{ps}{qr}. \quad (8)$$

Speciális esetként tekintsük a fenti két szituációt: ha ÉS-sel kapcsolt a bevásárlási kritérium, vagyis egymás hatását erősítik a kritériumok, akkor a férfiaknál a magas IQ a nőknél nagyobb hatást eredményez a jövedelemre:

$$\frac{p}{q} > \frac{r}{s},$$

ebből következően

$$\frac{ps}{qr} > 1,$$

vagyis $FEH_{JÖV=magas} > MEH$, a feltételes kapcsolat erősebb, mint a marginális. Vagyis marginális függetlenség esetén ez pozitív kapcsolat megjelenését jelenti a beválasztott alpopulációban.

Ha VAGY-gyal kapcsolt kritériumok, akkor egyik kritérium nem teljesülése esetén a másiknak megnő esélye, vagyis $r/s > p/q$, tehát $FEH < MEH$, vagyis marginális függetlenség esetén negatív lesz a feltételes kapcsolat. Ha gyakorlatilag kizárólagosak a kritériumok, vagyis az s 0 körüli érték, akkor a FEH is 0 körüli, erősen negatív kapcsolatot mutatva.

Ha a kritériumok egymástól függetlenül működnek, akkor $r/s = p/q$, s ebből az következik, hogy nem lép fel a Berkson-torzítás.

Fontos felhívni rá ismét a figyelmet, hogy bár a 7. ábrán a marginális függetlenséggel a Berkson-torzítás egy szélsőséges esetét ábrázoltam, a torzítás nem csak független beválasztási kritériumok esetén áll fenn, látható, hogy a FEH mindig módosul a MEH-hez képest.

A fentihez hasonló megfontolással belátható, hogy minden olyan alpopulációban áll-összefüggések rejlenek, ahol az al-populáció valamely, a populációban korrelálatlan beválogatási kritériumok alapján jött létre VAGY vagy ÉS-kapcsolás útján. Nyilván több kritérium valószínűsíti azt pl., hogy a kérdezett biztos pártválasztóként szerepel-e egy survey-ben, így a biztos pártválasztók között végzett, ezekkel a kritériumokkal kapcsolatos elemzések jó eséllyel torzítottak lehetnek. Például ha a szélsőjobboldali pártok szavazói és a vidékiek hajlamosabbak eltitkolni preferenciájukat, és ez a két (a teljes mintán pozitív kapcsolatot mutató) kritérium ÉS-sel van kapcsolva (vagyis egymás hatását erősíti), akkor a teljes mintán mért kapcsolatuknál a biztos szavazókon belül mért kapcsolat erősebb lesz. És nem csak a két változó kapcsolatát mérjük felül, hanem egymás meghatározó tényezőivel mért kapcsolatukat is: ha pl. a rasszizmus csak a szélsőjobboldali párt-preferenciával van a teljes mintán kapcsolatban, a vidéki volttal nem, a biztos szavazói mintán mégis azt fogjuk látni, hogy a vidékiek között több a rasszista.

Sőt, bármely survey-nél elmondható, hogy a válaszadási hajlandóságot több különböző kritérium teljesülése valószínűsítheti. Általában elmondható, hogy kevésbé szívesen vesznek részt felmérésekben a férfiak, a fiatalok, s a visszautasítási arány az iskolai végzettség szintjével nő. A munkanélküliek és nyugdíjasok nálunk felülreprezentáltak a lakossági mintákban, a gazdaságilag aktívak között jóval több a megtagadó, a nagyobb városokban és a fővárosban is magasabb a visszautasítók aránya (Pillók Péter, 2010). Ha ezek a faktorok interakcióban (akár ÉS-, akár VAGY-kapcsolásban) hatnak, akkor megjelenik a Berkson-torzítás. Arra vonatkozóan, hogy mely faktorok hatnak, és ezek közül melyek vannak interakcióban, általában csak közvetett ismereteink vannak, de ha címlista alapján választunk mintát, akkor néhány jellemző szerint ismerjük a sikertelen interjú adók megoszlását is. Pl. az OLEF2000 (Boros, Németh, Vitrai, 2002) felmérésben címlistát használtunk munkatársaimmal, és a sikertelen interjúk esetén is ismert volt a kérdezett neme, kora, lakhelye. Mindhárom tényező erős hatást gyakorolt a részvételre (jellemzően a budapestiek, a középkorúak/idősek és a nők között volt több a sikeres interjú), de e tényezők hatása

egymástól független volt (a $\frac{ps}{qr}$ értéke mindhárom változópár esetén két tizedesjegy pontossággal 1!), ezért e három tényező vonatkozásában nem lép fel a Berkson-torzítás.

6. Befejezés

A dolgozatban tárgyalt négy oksági megközelítés bizonyos szerzőknél konfliktusosan jelenik meg, s kizárólag egyik vagy másik elsőbbsége mellett érvelnek. Ezzel szemben én azt gondolom, hogy ezek a megközelítések nem kizáró dichotómiák. Ellenkezőleg, együttesen adnak egyfajta plurális oksági elméletet, hiszen az okság fogalmának mindegyikük egy-egy fontos aspektusát jeleníti meg. Egymással kompatibilissé tehetők, és - amennyiben más-más oksági bizonyítékot tárnak fel - egymás konklúzióit erősíthetik. Ami szembeállításukat kiváltja (lásd pl. *variable-sociology* vitát), az általában egy-egy létező, de sajnálatosan szélsőséges kutatási gyakorlat. Nyilván mindegyik megközelítéshez találhatunk olyan kutatási kérdéseket, amit az adott megközelítésben nem tudunk kezelni, de pl. talán nem lenne erős elvárás, hogy egy kauzális következtetés mindig hivatkozzon valószínűsíthető magyarázó mechanizmusokra is. Ugyanakkor az objektív, oksági bizonyítékot felmutatni képes kísérletes megközelítést is méltatlanul hiányzik sokszor a potenciális eszköztárból. Az igazán erős oksági következtetés mind a négy aspektusból alátámasztható lenne.²⁹

A dolgozat második felében (4. és 5. fejezet) ezeknek az aspektusoknak, s következetes használatuknak a hasznát próbáltam megmutatni. A tényellentétes modell szempontjából tárgyaltam a szociológiában fontos megfigyeléses vizsgálatok esetét, s megmutattam, hogy a megfigyeléses vizsgálatok alapproblémája, a *confounding* kézenfekvő módon tárgyalható e tényellentétes keretben. Az utolsó fejezetben az utóbbi években általam kutatott grafikus modelleket helyeztem el az okság kontextusában. A megfigyeléses vizsgálatok grafikus modellekkel megoldható alap-problémája (*adjustment-problem*) Pearl-féle megoldásának tárgyalása kapcsán kitértem a beválasztási torzításra, általánosítva az irodalomban leírt eseteket. A probléma gyakorlati fontosságára mutat rá a dolgozat befejező része, a torzítás megjelenésének lehetőségét tárgyalva a mindennapi survey-praxisban. Mivel a társadalomról alkotott ismereteink jó része survey-eken alapszik, ezt a torzítási lehetőséget érdemes lehet részletesebben is vizsgálni.

²⁹ Az epidemiológiában léteznek erre jó példák – például a dohányzás rákkeltő hatásával kapcsolatban. Az első megfigyeléses vizsgálatok a múlt század közepén még csak erős együttjárást tudtak kimutatni, ami elég robusztusnak bizonyult más faktorok bevonása után is (neo-humeiánus megközelítés). Ezt állatkísérletek is bizonyították (manipuláció), és világossá vált a leszokás kockázatsökkentő hatása is (tényellentétes). De a kauzális reláció bizonyításában a mögöttes, generáló folyamat megtalálásának, a karcinogén azonosításának volt döntő szerepe (mechanizmus-megközelítés).

7. Irodalom

- Andersson, Steen A., David Madigan and Michael D. Perlman (2001): Alternative Markov properties for chain graphs. *Scandinavian Journal of Statistics* 28(1):33–85.
- Angrist, Joshua D.; Krueger, Alan B. (1999): Empirical Strategies in Labor Economics. In: Ashenfelter, O.; Card, D. (szerk.): *Handbook of Labor Economics*. Vol. 3A. Elsevier; Amsterdam: 1999. p. 1277-1366.
- Arjas, Elja (2001): Causal analysis and statistics: a social sciences perspective. *European Sociological Review*, 17(1): 59-64.
- Bálint Lajos, Bozsonyi Károly (2012): Választói részvétel és véleménypolarizáció térfilterezett modelljei. In: Kmetty Zoltán, Koltai Júlia (szerk.): *Változó képletek, változatos perspektívák*. Tanulmánykötet Tardos Róbert 65. születésnapjára. Budapest: Háttér, 2012. pp. 251-274.
- Bearman, P. S., Moody, J., Stovel, K. (2002): Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, 110(1):44-91.
- Berk, R. A., Newton, P. (1985): Does arrest really deter wife battery? An effort to replicate the findings of the Minneapolis spouse abuse experiment. *American Sociological Review*, 50, 253-262.
- Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biometrics*. 1946;2:47–53.
- Bernert, C. (1983): The career of causal analysis in American sociology. *British Journal of Sociology*, 24(2): 230-254.
- Bickel, P.J., E.A. Hammel, J.W. O'Connell (1975). Sex Bias in Graduate Admissions: Data From Berkeley. *Science* 187 (4175): 398–404.
- Boros Julianna, Németh Renáta, Vitrai József (szerk.): Országos Lakossági Egészségfelmérés OLEF2000. Kutatási jelentés. Országos Epidemiológiai Központ, Budapest 2002.
- Brady, H. E. (2009): Causation and explanation in social science. In: J. M. Box-Steffensmeier, H. Brady és D. Collier (szerk.): *Oxford handbook of political methodology*, pp. 217-270. Oxford: Oxford University Press
- Bransen, J. (2001): Verstehen and erklären, the philosophy of. In: N. Smelser és P. Baltes (szerk.): *International Encyclopedia of the Social and Behavioral Sciences*, pp. 1665-1670. Oxford: Elsevier Science Ltd., 2001
- Blossfeld, Hans-Peter, Melinda Mills (2001): A Causal Approach to Interrelated Family Events: A Cross-national Comparison of Cohabitation, Nonmarital Conception, and Marriage. *Canadian Studies in Population*, 28(2):409.

- Boudon, R. (1976): Comment on Hauser's "Review of Education, Opportunity, and Social Inequality." *American Journal of Sociology*, 81:1175–87
- Cole, S.R., Platt, R.W., Schisterman, E.F., Chu, H., Westreich, D., Richardson, D., Poole, C. (2010): Illustrating bias due to conditioning on a collider. *International Journal of Epidemiology*, 39(2), 417-420.
- Collier, David, Henry E. Brady, Jason Seawright (2004): Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology. Chapter 13 in Henry E. Brady és David Collier (szerk.): *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman and Littlefield.
- Connors, A. F., Speroff, T., Dawson, N. V., et al. (1996): The effectiveness of right-heart catheterization in the initial care of critically ill patients. *Journal of the American Medical Association* 276: 889–97.
- Cook, T. D., Campbell, D. (1979): *Quasiexperimentation*. Chicago: Rand McNally.
- Cornfield J, et al. (1959): Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute*, 22(1):173–203.
- Cox, D. R. (1958): *The Planning of Experiments*. New York: Wiley.
- Cox, D.R., Wermuth, N. (1996). *Multivariate Dependencies – Models, Analysis and Interpretation*. London: Chapman & Hall.
- Drton, Mathias (2009): Discrete chain graph models. *Bernoulli* 15(3):736-753.
- Durkheim, Émile (1894): A szociológia módszertani szabályai. In: Durkheim, Émile: *A társadalmi tények magyarázatához*. Bp.: Közgazdasági és Jogi Kiadó, 1978. pp 141.
- E. Szabó László (2004): *A nyitott jövő problémája - Véletlen, kauzalitás és determinizmus a fizikában*, 6. fejezet: Kauzalitás (pp. 81-105). Budapest: Typotex.
- Elsenbroich, C (2012): Explanation in Agent-Based Modelling: Functions, Causality or Mechanisms? *Journal of Artificial Societies and Social Simulation*, 15(3):1.
- Elster, Jon (1983): *Explaining Technical Change*. Cambridge: Cambridge University Press.
- Elwert, Felix (2002): Incarceration, Unemployment and Eurosclerosis. Kutatási jelentés, Harvard University.
- Erdélyi Ágnes (2011): Az adekvát okozás problémája a történettudományban és a szociológiában. *Szociológiai Szemle*, 21(2): 18-31.
- Freedman, D. (2010): *Statistical Models and Causal Inference*. (Collier, D., Sekhon, J., Stark, J., eds.). New York: Cambridge University Press.
- Freedman, D.A, Pisani, R., Purves, R. (2005): *Statisztika*. Budapest, Typotex.
- Gárdos Judit (2014): *Előítéletek, romák, okok? Oksági modell egy magyarországi előítéletekutatásban*. Konferenciaelőadás, STS Műhelykonferencia, Budapest.

Goldthorpe, J. H. (2001): Causation, Statistics and Sociology. *European Sociological Review*, 17(1). 1–20.

Goldthorpe, J. H. (2005): Progress in Sociology: The Case of Social Mobility Research. In: S. Svallfors (ed.): *Analyzing Inequality: Life Chances and Social Mobility in Comparative Perspective*, pp. 56–82. Stanford: Stanford University Press.

Greenland, Sander; Robins, James Matthew (1986): Identifiability, Exchangeability and Confounding. *International Journal of Epidemiology*. 15(3): 413–419.

Greenland, Sander; Robins, James Matthew (2009): Identifiability, Exchangeability and Confounding revisited. *Epidemiologic Perspectives & Innovations*. 6(4): 1–9.

Hárs Ágnes, Simon Dávid (2013): Munkahelymegőrző válságprogramok bevezetése, eredményessége és hatásértékelése. *Külgazdaság*, 57. július–augusztus, 3–31.

Hauser, R. M. (1976): Review essay: on Boudon's model of social mobility. *American Journal of Sociology*, 81:911–28

Heckman, J. J. (2005): The scientific model of causality. *Sociological Methodology*, 35, 1–97.

Heckman J. J., Robb R. (1985): Alternative methods for evaluating the impact of interventions. In: Heckman J., Singer B. (szerk.): *Longitudinal Analysis of Labor Market Data*. Cambridge, UK: Cambridge Univ Press; pp. 156–245.

Hedström, P., Swedberg, R. (1996): Social Mechanisms. *Acta Sociologica*, 39(3):281–308.

Hedström, P. (2005): *Dissecting the Social: On the Principles of Analytical Sociology*, Cambridge: Cambridge University Press.

Hedström, P., Ylikoski, P. (2010): Causal Mechanisms in the Social Sciences. *Annual Review of Sociology*, 36:49–67.

Hernan, M., Hernandez-Diaz, S. and Robins, J. (2004). A structural approach to selection bias. *Epidemiology*, 15 615–625.

Holland, P. W. (1986): Statistics and Causal Inference (cikk, hozzászólások és viszontválasz). *Journal of the American Statistical Association*, 81(396): 945–960.

Huoranszki, Ferenc (2001): *Modern metafizika*. III. fejezet: Okság (pp. 51–73). Budapest: Osiris.

Jahoda, M. (1979): PFL: Hedgehog or Fox? In R. K. Merton, J. S. Coleman és P. H. Rossi (szerk.): *Qualitative and Quantitative Social Research: Papers in Honor of Paul F. Lazarsfeld*, pp. 3–9. New York: Free Press.

Julious, Steven A., Mark A. Mullee (1994): *Confounding* and Simpson's paradox. *British Medical Journal* 309 (6967): 1480–1481.

Kendall, P. L. és Lazarsfeld, P.F. (1950): Problems of Survey Analysis. In R. K. Merton és P. F. Lazarsfeld (szerk.): *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"*, Glencoe, Ill.: Free Press.

- Kertesi Gábor, Kézdi Gábor (2012): *A roma és nem roma tanulók tesztteredményei közti különbségekről és e különbségek okairól*. Budapesti Munkagazdaságtani Füzetek, 2012/5. Budapest: Budapesti Corvinus Egyetem.
- Kovács Balázs, Takács Károly (2003): Szimuláció a társadalomtudományokban. *Szociológiai Szemle*, 2003(3):27-49
- Lauritzen, Steffen L. (1996): *Graphical Models*. Oxford: Clarendon Press
- Lieberson, S. (1985): *Making It Count*. Berkeley: University of California Press.
- Marini, M. M., Singer, B. (1988): Causality in the social sciences. *Sociological Methodology*, 18(1): 347-409.
- Mauldon, J., Malvin, J., Stiles, J., Nicosia, N. és Seto, E. (2000): *Impact of California's Cal-Learn Demonstration Project: final report*. UC DATA Archive and Technical Assistance.
- Mayer-Schönberger, V., Cukier, K.: (2013): *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.
- Miettinen O. S., Cook, E. F. (1981): *Confounding: essence and detection*. *American Journal of Epidemiology*. 114: 593-603.
- Morgan, S. L., Winship, C. (2007): *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Analytical Methods for Social Research. New York: Cambridge University Press.
- Muthén, B. (1987): *Response to Freedman's critique of path analysis: Improve credibility by better methodological training*. *Journal of Educational Statistics*, 12, 178-184.
- Némethi Dénes (2000): A szociológia egy sikeres évszázad után. *Szociológiai Szemle*, 10(2): 3-16.
- Németh R., Rudas T. (2013a): On the application of discrete marginal graphical models. *Sociological Methodology*, 43:(1) pp. 70-100.
- Németh, R., Rudas T. (2013b): Discrete Graphical Models in Social Mobility Research: A Comparative Analysis of American, Czechoslovakian and Hungarian Mobility before the Collapse of State Socialism. *Bulletin de Methodologie Sociologique*, 118:(1) pp. 5-21.
- Németh Renáta (2013): Oksági következtetés a társadalomtudományi kutatásban. Előadás az MSZT konferenciáján, Budapest.
- Németh Renáta (2014): Oksági következtetés az empirikus szociológiai kutatásban. Kézirat, közlésre benyújtva.
- Németh Renáta, Rudas Tamás (2014): *Confounding in observational studies*. Kézirat.
- Oszlak, Oscar (1997): *The Argentine civil service: an unfinished search for identity*. A Civil Service Systems in Comparative Perspective c. konferencián elhangzott előadás, Indiana University, Bloomington, Indiana.
- Örkény Antal, Székelyi Mária (szerk.): *Deliberatív közvélemény-kutatás a magyarországi romák és nem romák viszonyáról*. Ulpius-ház, 2007.

- Pearl, J. (1995): Causal diagrams for empirical research. *Biometrika*, 82(4), pp. 662-710.
- Pearl, J. (2009): Causality: Models, Reasoning, and Inference. New York: Cambridge University Press.
- Pearl, J. (2014): Understanding Simpson's Paradox. *The American Statistician*, 68(1), 8-13.
- Pearson, Karl (2007[1911]): *The Grammar of Science*. Cosimo: New York.
- Pillók Péter (2010): *Az elhanyagolt tényező, avagy nem mintavételi hibák a kérdőíves adatfelvételekben*. Doktori disszertáció, ELTE TáTK, Budapest.
- Richardson, Thomas S. (2003): Markov Properties for Acyclic Directed Mixed Graphs. *Scandinavian Journal of Statistics* 30(1):145-157.
- Rosenbaum, P. R., Rubin, D. B. (1983): Assessing sensitivity to an unobserved binary covariate in an observational study with binary *outcome*. *Journal of the Royal Statistical Society* 45, 212-8.
- Rosenbaum, Paul R. (1984) The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society, Series A*, 147: 656-66.
- Rosenbaum, Paul R. (1999): Choice as an alternative to control in observational studies. *Statistical Science*, 14(3):259-278.
- Rubin, D. B. (1997): Estimating causal effects from large data sets using propensity scores. *Annals of International Medicine*, 127: 757-63
- Rudas Tamás (2010): Informative allocation and consistent treatment selection. *Statistical Methodology*, 7(3):323-337.
- Rudas, T., Bergsma, W., Németh, R. (2010): Marginal log-linear parameterization of conditional independence models. *Biometrika*
- Rudas Tamás, Németh Renáta (2013): Rejoinder: On the Application of Discrete Marginal Graphical Models *Sociological methodology*, 43(1): 130-132.
- Salmon, Wesley C. (1998): *Causality and Explanation*. Oxford: Oxford University Press.
- Schelling, T. C. (1978): *Micromotives and Macrobehavior*. New York: W.W. Norton
- Schild, M. (1999): Simpson's paradox and Cornfield's conditions. *American Statistical Association, Proceedings of the Section of Statistical Education*, 106-111.
- Sobel, M. E. (1995): Causal inference in the social and behavioral sciences. In: G. Arminger, C.C. Clogg és M.E. Sobel (szerk.): *A handbook for statistical modeling in the social and behavioral sciences*. New York: Plenum Press.
- Sørensen, A. B. (1998): Theoretical mechanisms and the empirical study of social processes. In: Hedström, P., Swedberg, R. (szerk.) *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge Univ. Press, pp. 238-66.

- Spellman, B. A., Mandel, D. R. (1999): When possibility informs reality: counterfactual thinking as a cue to causality. *Current Directions in Psychological Science*, 8: 120–3.
- Szántó Zoltán, Takács Károly (2012): Analitikus szociológia és hálózatelemzés. In: Kmetty Zoltán, Koltai Júlia (szerk.): *Változó képletek, változatos perspektívák. Tanulmánykötet Tardos Róbert 65. születésnapjára*. Budapest: Háttér Kiadó.
- Szántó Zoltán, SYI (Szakadát László) (2010): Fizikusok, bélyeggyűjtők, emberjárás-jelentők. *BUKSZ* 3., 201-213.
- Taubes, Gray (2001): The soft science of dietary fat. *Science*, 291(5513):2536-45.
- Tóth Gergely, Hajnóczky Tamás, Bozsonyi Károly (2013): A szegregációs folyamat egy sajátos modellje: A szegregációs folyamat történeti rekonstrukciója és annak matematikai modellje Tiszabő település esetén. In: Bacsák Dániel, Krámer Lili, Szabó Miklós (szerk.): *Kulcskérdések a társadalomtudományban*. Budapest: ELTE Társadalomtudományi Kar, 2013. pp. 201-222.
- Whittaker, Joe (1990): *Graphical Models in Applied Multivariate Statistics*. Chichester: Wiley.
- Willis, Robert J.; Rosen, Sherwin. (1979): Education and Self-Selection. *Journal of Political Economy*. 87:S7–36.
- von Wright, G. H. [1971] (1987): Magyarázat és megértés. In: Bertalan László (szerk.): *Magyarázat, megértés és előrejelzés*. Budapest: Tömegkommunikációs Kutatóközpont
- Xie, Yu (2013): Population heterogeneity and causal inference. *Proceedings of the National Academy of Sciences*, vol. 110, issue 16, pp. 6262-6268
- Zuberi, Tukufu (2001): *Thicker Than Blood: An Essay on how Racial Statistics Lie*. Minneapolis: University of Minnesota Press.